PRINCIPIOS DE MECANICA CUANTICA

PRINCIPIOS DE MECANICA CUANTICA

por

P. A. M. DIRAC
Profesor Lucasiano de Matemáticas
de la Universidad de Cambridge

EDICIONES ARIEL
Esplugues de Llobregat
Barcelona

Título original de la obra:

THE PRINCIPLES OF QUANTUM MECHANICS

Traducción de

ANTONIO MONTES

© 1958, Oxford University Press

© 1967, de la traducción castellana para España y América:
Ediciones Ariel, S. A. Barcelona

Impreso en España

Depósito legal: B. 18.229 - 1968

1968. - Ariel, S. A. - Avda. J. Antonio, 108. Esplugues de Llobregat - Barcelona

PRÓLOGO A LA CUARTA EDICIÓN

El cambio más importante respecto a la tercera edición es la nueva versión del capítulo que trata de la electrodinámica cuántica. La electrodinámica cuántica expuesta en la tercera edición describía el movimiento de partículas cargadas individuales, en estrecha analogía con la electrodinámica clásica. Era una forma de la teoría en que se conservaba necesariamente el número de partículas cargadas, y no podía generalizarse para permitir la variación del número de ellas.

Actualmente, en física de altas energías se da con frecuencia la creación y aniquilación de partículas cargadas. Por consiguiente, una electrodinámica cuántica que exija la conservación del número de partículas cargadas está lejos de la realidad física. Así pues, la hemos substituido por una electrodinámica cuántica que incluye la creación y aniquilación de pares electrón-positrón. Ello nos obliga a abandonar toda analogía estrecha con la teoría clásica de los electrones, pero nos proporciona una descripción más ajustada de la naturaleza. Parece pues que el concepto clásico de electrón deja de ser un modelo válido en física, salvo para algunas teorías elementales que se limiten a considerar fenómenos de baja energía.

P. A. M. D.

St. John's College, Cambridge
11 de mayo de 1957

DEL PRÓLOGO A LA PRIMERA EDICIÓN

Los métodos de trabajo de la física teórica han sufrido un profundo cambio en el presente siglo. La tradición clásica consideraba el mundo como la asociación de objetos observables (partículas, fluidos, campos, etc.) que se mueven según leyes de fuerza fijas, pudiéndose, pues, formar una imagen mental de todo el esquema en el espacio y el tiempo. Ello conducía a una física cuyo método consistía en hacer hipótesis sobre el mecanismo y las fuerzas que relacionan dichos objetos observables, que explicaran su comportamiento de la forma más sencilla posible. En los últimos tiempos se ha ido haciendo cada vez más evidente que la naturaleza actúa en un plano distinto. Sus leyes fundamentales no rigen el mundo directamente tal como éste aparece en nuestra imagen mental, sino que actúan sobre un substrato del que no podemos formarnos ninguna imagen mental sin cometer desatinos. La formulación de dichas leyes exige el empleo de las matemáticas de transformaciones. Los elementos importantes del mundo físico aparecen como los invariantes de dichas transformaciones (o más en general los casi-invariantes, o cantidades que se transforman de modo sencillo). Las cosas que conocemos de modo inmediato son las relaciones de estos casi-invariantes con un cierto sistema de referencia, que generalmente elegimos de forma que se introduzcan simplificaciones particulares sin importancia desde el punto de vista de la teoría general.

El auge de la teoría de transformaciones, aplicada primero a la relatividad y después a la teoría cuántica, constituye el fundamento de los nuevos métodos de la física teórica. Otra característica del desarrollo de la teoría es la obtención de ecuaciones que sean invariantes frente a transformaciones cada vez más generales. Este modo de proceder es muy satisfactorio desde el punto de vista filosófico, pues reconoce el verdadero papel del observador, permitiendo englobar las reglas que se deducen de sus observaciones particulares en un esquema único, y mostrando que la naturaleza no se comporta de modo arbitrario; pero a cambio dificulta la comprensión para el estudioso de la física. Las nuevas teorías, si se consideran independientemente de su estructura matemática, están constituidas por conceptos físicos que no pueden expresarse mediante términos o elementos conocidos previamente por el estudiante, y que ni siquiera pueden explicitarse adecuadamente con palabras. Al igual que los conceptos fundamentales (como por ejemplo, proximidad, identidad) que cada uno ha de aprender desde que nace, los nuevos conceptos de la física sólo pueden dominarse familiarizándose con sus propiedades y usos.

Desde el punto de vista matemático, la comprensión de las nuevas teorías no presenta ninguna dificultad, ya que las matemáticas necesarias (por lo menos las que se necesitan para el desarrollo de la física hasta nuestros días) no son esencialmente distintas de las que se empleaban anteriormente. Las matemáticas son el instrumento especialmente indicado para tratar conceptos abstractos de cualquier clase, y en este campo su poder no tiene límites. Por esta razón todo libro sobre la nueva física, que no sea puramente una descripción de trabajos experimentales, ha de ser esencialmente matemático. Sin embargo, las matemáticas no son más que un instrumento y deberíamos aprender a apropiarnos de las ideas físicas sin referencia a su forma matemática. En este libro hemos procurado poner en primer plano el aspecto físico, comenzando por un capítulo enteramente físico y procurando examinar en lo posible el significado físico que encierra el formalismo en los capítulos siguientes. Se necesita gran cantidad de conocimientos teóricos para poder resolver problemas de valor práctico, pero este hecho es una consecuencia inevitable del importante papel que tiene la teoría de transformaciones y que está destinado a acentuarse en la física teórica del futuro.

Respecto al formalismo matemático en que puede presentarse la teoría, el autor ha de decidirse desde el principio entre dos métodos. Uno es el método simbólico, que considera directamente y de forma abstracta las cantidades de importancia fundamental (invariantes, etc., de las transformaciones) y el otro se basa en el empleo de coordenadas o representaciones y trata con conjuntos de números que corresponden a dichas cantidades. Por regla general siempre se ha utilizado el segundo método para presentar la mecánica cuántica (de hecho se ha utilizado prácticamente siempre a excepción del libro de Weyl titulado Gruppentheorie und Quantenmechanik). *Dicho método se conoce con los nombres de 'mecánica ondulatoria' y 'mecánica de las matrices' según a cual de los elementos físicos, estados o variables dinámicas del sistema, se le conceda mayor importancia. Tiene la ventaja de que las matemáticas que exige son más familiares para el estudiante medio y además es el método histórico.*

Sin embargo, el método simbólico parece atacar con más profundidad la naturaleza de las cosas. Nos permite expresar las leyes físicas de una forma clara y concisa, y probablemente se irá utilizando cada vez más en el futuro a medida que vaya siendo mejor conocido y se vayan desarrollando las matemáticas especiales que emplea. Por esta razón hemos elegido el método simbólico, introduciendo sólo después las representaciones como ayuda para los cálculos prácticos. Ello ha exigido romper completamente con la línea de desarrollo histórica, pero nos permite el acceso a las nuevas ideas de la forma más directa posible.

P. A. M. D.

St. John's College, Cambridge
29 de mayo de 1930

ÍNDICE

I. El principio de superposición 15

1. Necesidad de una teoría cuántica 15
2. Polarización de fotones 18
3. Interferencia de fotones 21
4. Superposición e incertidumbre 23
5. Formulación matemática del principio de superposición . 27
6. Vectores bra y ket 30

II. Variables dinámicas y observables 35

7. Operadores lineales 35
8. Relaciones conjugadas 38
9. Autovalores y autovectores 41
10. Observables 46
11. Funciones de observables 52
12. Interpretación física general 56
13. Conmutabilidad y compatibilidad 60

III. Representaciones 64

14. Vectores básicos 64
15. La función δ 69
16. Propiedades de los vectores básicos 72
17. Representación de operadores lineales 78
18. Amplitudes de probabilidad 83
19. Teoremos sobre funciones de observables 87
20. Nuevas formas de notación 89

IV. Condiciones cuánticas 95

21. Corchetes de Poisson 95
22. Representación de Schrödinger 99
23. La representación de momentos 106
24. El principio de incertidumbre de Heisenberg . . . 109
25. Operadores de traslación 110
26. Transformaciones unitarias 114

V. Las ecuaciones del movimiento 120
27. Ecuaciones del movimiento en imagen de Schrödinger . 120
28. Ecuaciones del movimiento en imagen de Heisenberg . . 123
29. Estados estacionarios 128
30. La partícula libre 130
31. Movimiento de paquetes de ondas 133
32. El principio de acción 137
33. El conjunto de Gibbs 143

VI. Aplicaciones elementales 148
34. El oscilador armónico 148
35. Momento angular 152
36. Propiedades del momento angular 156
37. El spin del electrón 161
38. Movimiento en un campo de fuerzas central . . . 165
39. Niveles de energía del átomo de hidrógeno 169
40. Reglas de selección 172
41. El efecto de Zeeman en el átomo de hidrógeno . . . 177

VII. Teoría de perturbaciones 180
42. Generalidades 180
43. Variación de los niveles de energía producida por una perturbación 181
44. La perturbación considerada como causa de transiciones . 185
45. Aplicación a la radiación 189
46. Transiciones producidas por una perturbación independiente del tiempo 191
47. El efecto Zeeman anómalo 194

VIII. Problemas de colisión 199
48. Observaciones generales 199
49. Coeficiente de dispersión 202
50. Solución en la representación de momentos 207
51. Dispersión con absorción y reemisión 213
52. Dispersión de resonancia 215
53. Emisión y absorción 218

IX. Sistemas que contienen partículas idénticas 221
54. Espacios simétricos y antisimétricos 221
55. Las permutaciones como variables dinámicas 225
56. Las permutaciones como constantes del movimiento . . 227
57. Determinación de los niveles de energía 230
58. Aplicación a electrones 233

X. Teoría de la radiación 239
59. Conjuntos de bosones 239
60. Relación entre bosones y osciladores 242
61. Emisión y absorción de bosones 247
62. Aplicación a fotones 250
63. La energía de interacción de los fotones con un átomo . 254
64. Emisión, absorción y dispersión de la radiación . . . 259
65. Conjuntos de fermiones 263

XI. Teoría relativista del electrón 268
66. Teoría relativista de una partícula 268
67. La ecuación de onda para el electrón 270
68. Invariancia bajo una transformación de Lorentz . . . 273
69. Movimiento de un electrón libre 276
70. Existencia del spin 279
71. Transformación a variables polares 283
72. Estructura fina de los niveles de energía del hidrógeno . 285
73. Teoría del positrón 289

XII. Electrodinámica cuántica 292
74. El campo electromagnético en ausencia de materia . . . 292
75. Expresión relativista de las condiciones cuánticas . . 296
76. Variables dinámicas de Schrödinger 299
77. Condiciones suplementarias 304
78. Electrones y positrones aislados 308
79. La interacción 315
80. Variables físicas 319
81. Dificultades de la teoría 324

NO

KA Lifshitz, Pitaevskii y Bereteskii
Teoría cuántica relativista.

— * Berestetskii
Quantum electrodynamics

— * R.P. Feynman Quantum electrodynamics

I

EL PRINCIPIO DE SUPERPOSICIÓN

1. Necesidad de una teoría cuántica

La mecánica clásica se ha ido desarrollando progresivamente desde los tiempos de Newton y se ha aplicado a un conjunto de sistemas dinámicos cada vez más amplio, del que forma parte el campo electromagnético en interacción con la materia. Las ideas fundamentales y las leyes que rigen su aplicación constituyen un esquema tan sencillo y elegante, que parece imposible modificarlo seriamente sin destruir todas sus atractivas características. Sin embargo, se ha conseguido construir un nuevo esquema, llamado mecánica cuántica, más adecuado para la descripción de los fenómenos de escala atómica, y que es en ciertos aspectos más elegante y satisfactorio que el esquema clásico. Esto ha sido posible gracias a que los cambios introducidos por la nueva teoría son de carácter muy profundo y no simples modificaciones que destruirían las características de la teoría clásica que le confieren su armonía, y así ha resultado que todas estas características han podido ser incorporadas al nuevo esquema.

La necesidad de elegir un camino distinto al de la mecánica clásica viene exigida por los hechos experimentales. En primer lugar, las fuerzas conocidas en electrodinámica clásica son inadecuadas para justificar la notable estabilidad de átomos y moléculas, necesaria para poder explicar que las substancias tengan propiedades físicas y químicas definidas. La introducción de nuevas fuerzas hipotéticas no podría salvar la situación, ya que existen principios generales de la mecánica clásica, válidos para cualquier tipo de fuerzas, que nos conducen a resultados en completo desacuerdo con la observación. Por ejemplo, si el estado de equilibrio de un sistema atómico está alterado de algún modo y se le abandona a sí mismo en estas condiciones, debería comenzar a oscilar, y sus oscilaciones tendrían que ponerse de manifiesto en el campo electromagnético que se emite, de modo que sus frecuencias habrían de ser observables con un espectroscopio. Cualesquiera que fuesen las fuerzas que rigen este equilibrio, debería ser posible incluir las diversas frecuencias en un cuadro que comprendiera ciertas frecuencias fundamentales y sus armónicos. Esto no corresponde a lo que se observa.

En su lugar aparece una nueva e imprevista relación entre las frecuencias, llamada ley de combinación de Ritz de la espectroscopia, según la cual todas las frecuencias pueden ser expresadas como diferencias entre ciertos términos, siendo el número de términos muy inferior al de frecuencias. Esta ley es completamente incomprensible desde el punto de vista clásico.

Se puede intentar salvar la dificultad sin apartarse de la mecánica clásica, suponiendo que cada una de las frecuencias observadas con medidas espectroscópicas es una frecuencia fundamental inherente a su propio grado de libertad, siendo las leyes de fuerza tales que no puedan darse los armónicos correspondientes. Sin embargo, una teoría como ésta no sería válida, aparte de no dar explicación alguna de la ley de combinación, ya que entraría en conflicto inmediato con la evidencia experimental sobre calores específicos. La mecánica estadística clásica nos permite establecer una relación general entre el número total de grados de libertad de un conjunto de sistemas vibrantes y su calor específico. Si se hiciera la hipótesis de que todas las frecuencias espectroscópicas de un átomo corresponden a distintos grados de libertad, se obtendría un calor específico para todas las sustancias mucho mayor que el observado. De hecho, los calores específicos observados a temperaturas ordinarias vienen dados con bastante exactitud por una teoría que considere el movimiento de cada átomo como el de un todo sin atribuirle ningún movimiento interno.

Esto nos lleva a un nuevo antagonismo entre la mecánica clásica y los resultados experimentales. Debe existir necesariamente algún movimiento interno en los átomos para poder explicar su espectro; pero sus grados de libertad internos, por razones incomprensibles clásicamente, no contribuyen al calor específico. Otro conflicto similar se encuentra a propósito de la energía de oscilación del campo electromagnético en el vacío. La mecánica clásica da un valor infinito para el calor específico correspondiente a esta energía, pero se ha podido comprobar que es perfectamente finito. La conclusión general de los resultados experimentales es que las oscilaciones de gran frecuencia no contribuyen en su valor clásico al calor específico.

Un ejemplo más de la limitación de la mecánica clásica lo constituye el comportamiento de la luz. Tenemos, por un lado, los fenómenos de interferencias y de difracción, que sólo pueden ser explicados mediante una teoría ondulatoria; por el otro, fenómenos tales como la emisión fotoeléctrica y la dispersión (scattering), de electrones libres, que indican que la luz está compuesta por pequeñas partículas. Dichas partículas, llamadas fotones, tienen cada una una energía y un momento definidos, que dependen de la frecuencia de la luz, y aparecen con una existencia tan real como la de los electrones o cualesquiera otras partículas conocidas en física. No se observa nunca una fracción de fotón.

Los experimentos han demostrado que este extraño comportamiento no es peculiar de la luz, sino que es completamente general. Todas las partículas materiales tienen propiedades ondulatorias, que se ponen de manifiesto

en condiciones adecuadas. Éste es un ejemplo muy sorprendente y característico de los fallos de la mecánica clásica, que no radican simplemente en una inexactitud de sus leyes del movimiento, sino en *una insuficiencia de sus conceptos para proporcionarnos una descripción de los fenómenos atómicos.*

La necesidad de apartarse de las ideas clásicas al intentar dar una explicación de la estructura elemental de la materia se deriva no sólo de los hechos establecidos experimentalmente, sino también de razones filosóficas generales. En una interpretación clásica de la constitución de la materia, se supondría que ésta está compuesta por un gran número de pequeñas partes, y se postularían leyes del comportamiento de dichas partes, de las cuales se pudieran deducir las leyes de la materia agrupada. Sin embargo, no sería una explicación completa, pues habría quedado sin considerar la estructura y estabilidad de las partes constituyentes. Para responder a esta cuestión se haría necesario postular que cada una de las partes constituyentes está a su vez compuesta de pequeñas partes, en razón de las cuales debe explicarse su comportamiento. Evidentemente no hay límite para este proceso, de tal forma que es imposible llegar por este procedimiento a la estructura elemental de la materia. Mientras *grande* y *pequeño* sean conceptos meramente relativos, no conduce a nada explicar lo grande en función de lo pequeño. Por tanto, es necesario que modifiquemos las ideas clásicas de forma que el tamaño adquiera un carácter absoluto.

En este punto es importante recordar que a la ciencia solamente le incumben los objetos observables y que únicamente podemos observar un objeto si interacciona con alguna influencia externa. Todo acto de observación va acompañado necesariamente de una alteración del objeto observado. Podemos decir que un objeto es grande cuando la alteración que acompaña a nuestra observación de él pueda ser despreciada, y pequeño cuando no pueda serlo. Esta definición está plenamente de acuerdo con el significado corriente de las expresiones grande y pequeño.

Normalmente se supone que, procediendo con cuidado, podemos reducir la alteración que acompaña a nuestra observación a una amplitud tan pequeña como queramos. En este caso, los conceptos grande y pequeño son meramente relativos y se refieren tanto a la precisión de nuestros medios de observación como al objeto que se describe. Para poder dar un significado absoluto a la magnitud, tal como se requiere en toda teoría de la estructura elemental de la materia, hemos de suponer que *existe un límite de la precisión de nuestro poder de observación y de la magnitud de la alteración que le acompaña — límite que es inherente a la naturaleza de las cosas y que es imposible superar aunque se perfeccionen las técnicas o se aumente la habilidad práctica del observador*. Si el objeto que consideramos es tal que la alteración límite inevitable se puede despreciar, decimos que el objeto es grande en sentido absoluto y podremos aplicarle la mecánica clásica. Si, en cambio, dicha alteración no es despreciable, el objeto es

pequeño en sentido absoluto y será necesaria una nueva teoría para dar cuenta de él.

Como consecuencia de la discusión precedente debemos revisar nuestras ideas sobre la causalidad. La relación causal se aplica exclusivamente a sistemas que no hayan sido alterados. Pero si el sistema es pequeño, nos es imposible observarlo sin producir en él una alteración considerable, y por lo tanto, no podemos esperar encontrar ninguna relación causal entre los resultados de nuestras medidas. Supondremos que sigue existiendo causalidad para los sistemas mientras no se les altere, y las ecuaciones que estableceremos para describir los sistemas no alterados serán ecuaciones diferenciales que expresarán una relación causal entre condiciones en un instante y condiciones en otro instante posterior. Estas ecuaciones estarán en estrecha correspondencia con las ecuaciones de la mecánica clásica, pero sólo estarán indirectamente relacionadas con los resultados de las observaciones. Existe, pues, una indeterminación inevitable en el cálculo de los resultados observables, y la teoría no nos permite calcular, en general, más que la probabilidad de obtener un resultado particular al hacer la observación.

2. *Polarización de fotones*

La discusión de la sección precedente acerca del límite de la precisión con que pueden ser realizadas las observaciones y la indeterminación consiguiente en los resultados de dichas observaciones no nos proporciona ninguna base cuantitativa para la construcción de la mecánica cuántica. Para este fin necesitamos un nuevo conjunto adecuado de leyes de la naturaleza. Una de las fundamentales y más eficaces es el *principio de superposición de los estados*. Llegaremos a la formulación general de este principio a través de algunos casos particulares, y tomaremos como primer ejemplo el que nos proporciona la polarización de la luz.

Se sabe experimentalmente que cuando se utiliza luz de polarización rectilínea para expulsar fotoelectrones, existe una dirección privilegiada de la emisión electrónica. Por tanto, las propiedades de polarización de la luz están íntimamente relacionadas con sus propiedades corpusculares y así se debe atribuir una polarización a los fotones. Así pues, debemos considerar que un haz de luz polarizada rectilíneamente en una cierta dirección está constituido por fotones, cada uno de los cuales está polarizado rectilíneamente en dicha dirección, y que asimismo un haz de luz polarizada circularmente está constituido por fotones cada uno de ellos polarizado circularmente. Cada fotón está en un cierto estado, llamado *estado de polarización*. El problema que hemos de considerar es cómo adaptar estas ideas a los fenómenos conocidos de la separación de la luz en componentes polarizadas y de la recombinación de estas componentes.

Elijamos un caso concreto. Supongamos que tenemos un haz de luz que

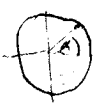
atraviesa un cristal de turmalina, que tiene la propiedad de dejar pasar únicamente luz polarizada según un plano perpendicular a su eje óptico. La electrodinámica clásica explica lo que ocurre para cualquier polarización del haz incidente. Si dicho haz está polarizado según un plano perpendicular al eje óptico, todo él atravesará el cristal; si lo está según un plano paralelo al eje, no lo atravesará en absoluto; mientras que si está polarizado según una dirección que forma un ángulo α con el eje, atravesará una fracción $\text{sen}^2\alpha$ de él. ¿Cómo podemos interpretar estos resultados sobre la base fotónica?

Un haz polarizado rectilíneamente en una cierta dirección debe imaginarse constituido por fotones cada uno de ellos polarizado en esa dirección. Esta imagen no presenta ninguna dificultad en los casos en que el haz incidente está polarizado según un plano perpendicular o paralelo al eje óptico. No tenemos más que suponer que cada fotón polarizado perpendicularmente al eje atraviesa el cristal sin obstáculo y sin sufrir cambio alguno, mientras que si está polarizado paralelamente al eje es frenado y absorbido. La dificultad aparece, en cambio, en el caso de polarización oblicua y no es evidente lo que le debe ocurrir a uno de tales fotones cuando alcanza la turmalina.

La pregunta de qué le ocurre a un determinado fotón bajo unas condiciones definidas no tiene un significado demasiado preciso. Para precisarlo debemos imaginar algún experimento que tenga relación con ella y preguntarnos cuál será el resultado del mismo. Sólo tienen significado real las preguntas sobre los resultados de experimentos y son éstas las únicas preguntas que debe contestar la física teórica.

En nuestro ejemplo dicho experimento consiste en utilizar un haz incidente constituido por un único fotón y observar qué es lo que aparece al otro lado del cristal. Según la mecánica cuántica el resultado de este experimento es que al otro lado encontraremos unas veces un fotón entero de energía igual a la del fotón incidente, y otras no aparecerá nada. Cuando aparezca un fotón entero, estará polarizado perpendicularmente al eje óptico. En ningún caso encontraremos al otro lado sólo una parte de fotón. Si repetimos el experimento un gran número de veces encontraremos que aparece un fotón por el otro lado en una proporción $\text{sen}^2\alpha$ del número total de veces. Podemos decir que el fotón tiene una probabilidad $\text{sen}^2\alpha$ de pasar a través de la turmalina y aparecer por el otro lado polarizado según un plano perpendicular al eje, y una probabilidad $\cos^2\alpha$ de ser absorbido. Tales valores de la probabilidad conducen a los resultados clásicos correctos en el caso de un haz incidente con un gran número de fotones.

Por este procedimiento conservamos en todos los casos la individualidad del fotón. Pero ello sólo es posible si abandonamos el determinismo de la teoría clásica. El resultado de un experimento no está determinado por condiciones controlables por el experimentador como ocurriría de acuerdo con las ideas clásicas. Lo máximo que podemos predecir es un conjunto



$n \propto \text{sen}^2\alpha$

de resultados posibles, y la probabilidad de que ocurra cada uno de ellos.

La discusión precedente sobre el resultado de un experimento con un único fotón polarizado oblicuamente que incide sobre un cristal de turmalina contesta a todas las preguntas que se pueden formular legítimamente sobre lo que le ocurre a un fotón bajo esas condiciones. Preguntas tales como qué es lo que decide si el fotón atraviesa o no el cristal y cómo cambia su dirección de polarización cuando lo atraviesa, no pueden ser investigadas mediante experimentos y deben considerarse fuera del dominio de la ciencia. No obstante, para relacionar los resultados de este experimento con los de otros experimentos que se pueden realizar con fotones y englobarlos todos dentro de un esquema general, es necesaria una descripción ulterior. No debe considerarse dicha descripción como un intento de responder a problemas que están fuera del dominio de la ciencia, sino como una ayuda para la formulación de leyes que expresen de forma concisa los resultados de un gran número de experimentos.

La descripción ulterior que nos proporciona la mecánica cuántica es la siguiente. Se supone que un fotón polarizado según un plano oblicuo al eje óptico puede considerarse que está en parte en el estado de polarización paralelo al eje y en parte en el estado de polarización perpendicular al eje. El estado de polarización oblicua puede ser considerado como el resultado de un cierto proceso de superposición aplicado a los dos estados de polarización paralela y perpendicular. Ello implica una cierta relación entre los diversos estados de polarización similar a la de la óptica clásica para haces polarizados, pero que ahora no se aplica a los haces sino a los estados de polarización de cada fotón particular. Esta relación nos permite resolver o expresar cualquier estado de polarización como una superposición de un par arbitrario de estados de polarización perpendiculares entre sí.

Cuando hacemos chocar un fotón con un cristal de turmalina, le sometemos a una observación. Observamos si está polarizado bien paralela o bien perperdicularmente al eje óptico. El efecto de esta observación es obligar al fotón a que esté enteramente en el estado de polarización paralela o enteramente en el estado de polarización perpendicular. Se ve obligado a dar un salto brusco de estar parcialmente en cada uno de dichos estados a estar enteramente en uno u otro de ellos. No se puede predecir a cuál de los dos estados pasará, y únicamente existe una ley que nos da cada una de las dos probabilidades. Si salta al estado de polarización paralela será absorbido y si salta al de polarización perpendicular atravesará el cristal y aparecerá por el otro lado conservando dicho estado de polarización.

3. *Interferencia de fotones*

En esta sección vamos a considerar otro ejemplo de superposición. También utilizaremos fotones, pero ahora nos interesaremos por su posición en el espacio y por su momento, en lugar de interesarnos por su polarización. Si disponemos de un haz de luz estrictamente monocromática, conocemos algo acerca de la posición y momento de los fotones asociados. Sabemos que cada uno de ellos está localizado en algún lugar de la región del espacio por donde pasa el haz, y que tiene un momento según la dirección del haz cuya magnitud viene dada en función de la frecuencia de éste por la ley de Einstein del efecto fotoeléctrico — momento igual a frecuencia multiplicada por una constante universal. Cuando tengamos una información como ésta sobre la posición y el momento de un fotón diremos que está en un *estado de traslación* definido.

Expliquemos la descripción que da la mecánica cuántica de la interferencia de fotones. Elijamos un experimento concreto de interferencia. Supongamos que un haz de luz ha pasado a través de un interferómetro, y que se ha dividido en dos componentes que se hacen interferir a continuación. Como en la sección precedente, podemos elegir el caso de un haz constituido por un único fotón, y preguntarnos qué le ocurrirá cuando atraviese el aparato. Así se pondrá claramente de manifiesto el conflicto entre las teorías ondulatoria y corpuscular de la luz.

De acuerdo con la descripción que hemos dado en el caso de la polarización, tenemos que considerar que el fotón está parcialmente en cada una de las dos componentes en que se ha dividido el haz incidente. Diremos que el fotón se halla en un estado de traslación dado por la superposición de los dos estados de traslación asociados a las dos componentes. Esto nos induce a generalizar el término 'estado de traslación' aplicado a fotones. Para que un fotón esté en un estado de traslación definido no es necesario que esté asociado a un único haz de luz, sino que puede estarlo a dos o más haces, que son las componentes en que se ha dividido el haz inicial.* En la teoría matemática rigurosa, cada estado de traslación está asociado a una de las funciones de onda de la óptica ondulatoria clásica, que puede representar un único haz o bien dos o más haces en los que se haya dividido el haz inicial. Los estados de traslación se pueden superponer de forma análoga a las funciones de onda.

Veamos qué ocurre cuando determinamos la energía en una de las componentes. El resultado de tal determinación ha de ser o bien que encontremos el fotón o que no encontremos nada. Así pues, el fotón debe cambiar bruscamente de estar parcialmente en uno de los haces y parcialmente en el

---

* El hecho de que la idea de superposición nos obligue a generalizar el significado de los estados de traslación y no, en cambio, el de los estados de polarización de la sección anterior es accidental y no posee ningún significado teórico fundamental.

otro a estar totalmente en uno de ellos. Este cambio brusco es debido a la alteración del estado de traslación del fotón que necesariamente lleva consigo toda observación. Es imposible predecir en cuál de los dos haces encontraremos el fotón. Únicamente podemos calcular la probabilidad de cada resultado basándonos en la distribución previa de fotones en cada haz.

Se podría llevar a cabo la medición de energía sin destruir el haz componente, por ejemplo, reflejándolo sobre un espejo móvil y observando su retroceso. Nuestra descripción de los fotones permite inferir que, *después* de tal medición de energía, no sería posible llevar a cabo ningún efecto de interferencia entre las dos componentes. Mientras el fotón está parcialmente en cada uno de los haces pueden producirse interferencias al superponerse ambos haces, pero dicha posibilidad desaparece cuando, mediante una observación, obligamos al fotón a estar enteramente en uno de los haces. En el último caso el otro haz ya no interviene para nada en la descripción del fotón, de manera que debe considerársele íntegramente en el primero de ellos como en el caso ordinario, a efectos de cualquier experimento que pueda realizarse a continuación con él.

De este modo la mecánica cuántica permite reconciliar las teorías ondulatoria y corpuscular de la luz. El punto fundamental es la asociación de cada estado de traslación de un fotón a una de las funciones de onda de la óptica ondulatoria ordinaria. La naturaleza de esta asociación no puede imaginarse sobre la base de la mecánica clásica; es algo completamente nuevo. Sería totalmente incorrecto imaginar que el fotón y su onda asociada están en interacción y suponer que dicha interacción es análoga a las que se dan en mecánica clásica entre partículas y ondas. La interpretación de dicha asociación sólo puede ser de naturaleza estadística, en la que la función de onda nos da información acerca de la probabilidad de encontrar el fotón en cualquier lugar determinado al realizar una observación de su posición.

Algún tiempo antes del descubrimiento de la mecánica cuántica, los científicos se dieron cuenta de que la relación entre ondas luminosas y fotones debía ser de naturaleza estadística, pero lo que no comprendieron claramente es que la función de onda dé información sobre la probabilidad de que *un* fotón se encuentre en un determinado lugar y no sobre el número probable de fotones en dicho lugar. La importancia de esta distinción se puede poner de manifiesto del modo siguiente. Sea un haz de luz constituido por un gran número de fotones dividido en dos componentes de igual intensidad. En la hipótesis de que la intensidad del haz esté relacionada con el número probable de fotones en él, encontraremos la mitad de fotones en cada una de las dos componentes. Si hacemos que ambas componentes interfieran, deberíamos admitir que un fotón de una componente es capaz de interferir con uno de la otra. Algunas veces estos dos fotones tendrían que aniquilarse uno con otro, y otras veces deberían dar lugar a cuatro fotones. Esto contradice el principio de conservación de la energía.

La nueva teoría, que relaciona la función de onda con la probabilidad de un único fotón, supera la dificultad al suponer que el fotón está parcialmente en cada una de las dos componentes. Cada fotón interfiere únicamente consigo mismo. Nunca pueden ocurrir fenómenos de interferencias entre dos fotones distintos.

La asociación de partículas y ondas que acabamos de exponer no se limita al caso de la luz; según la nueva teoría es de validez general. Todo tipo de partículas tiene una onda asociada en esta forma, y recíprocamente, a todo movimiento ondulatorio se le asocian partículas. Por tanto, todas las partículas son susceptibles de experimentar fenómenos de interferencias y todas las ondas tienen su energía en forma de cuantos. La razón de que estos fenómenos generales no sean más manifiestos se debe a la ley de proporcionalidad entre la masa o la energía y la frecuencia de las ondas, con un coeficiente de proporcionalidad tal, que para ondas de frecuencias ordinarias el cuanto asociado es extremadamente pequeño, mientras que para partículas como los electrones, o incluso la luz, la frecuencia de la onda asociada es tan grande que no es fácil verificar con ellas fenómenos de interferencias.

4. *Superposición e incertidumbre*

Probablemente el lector se encontrará insatisfecho con el intento realizado en las dos secciones precedentes, para compaginar la existencia de los fotones y la teoría clásica de la luz. Puede alegar que se ha introducido una idea muy singular — la posibilidad de que un fotón esté parcialmente en cada uno de dos estados de polarización o en cada uno de dos haces separados — y que a pesar de ello no se ha dado ninguna imagen satisfactoria del proceso elemental de un fotón. Puede añadir que dicha idea no proporciona más información de la que se podría obtener mediante una consideración elemental donde los fotones estuvieran conducidos por ondas de forma vagamente definida. ¿Cuál es, pues, la finalidad de esta idea singular?

Como respuesta a la primera crítica debe advertirse que el principal objeto de la ciencia física no es dar imágenes sino formular las leyes que rigen los fenómenos y su aplicación para descubrir nuevos fenómenos. Si existe una imagen tanto mejor; pero el que exista o no es un hecho de importancia secundaria. En el caso de los fenómenos atómicos, no puede esperarse que exista ninguna 'imagen' en el sentido habitual de la palabra, que viene a significar un modelo que funcione esencialmente según la mecánica clásica. Sin embargo, se puede generalizar el significado de la palabra 'imagen' de modo que incluya *cualquier forma de considerar las leyes fundamentales que haga evidente su coherencia*. Con esta generalización se puede adquirir gradualmente una imagen de los fenómenos atómicos al irse familiarizando con las leyes de la mecánica cuántica.

Respecto a la segunda crítica debe tenerse en cuenta que para muchos experimentos sencillos con la luz sería suficiente, a fin de explicar los resultados, una teoría elemental en la que ondas y partículas estuviesen relacionadas de una vaga forma estadística. La mecánica cuántica no aporta ninguna otra información sobre dichos experimentos. Sin embargo, en la gran mayoría de experimentos, las condiciones son demasiado complejas para que sea aplicable una teoría tan elemental, y es necesario algún esquema más elaborado como el de la mecánica cuántica. El método de descripción que da la mecánica cuántica para los casos más complejos también se puede aplicar a los casos sencillos, y si bien para ellos no es estrictamente necesario, su estudio en estos casos constituye una introducción apropiada para el estudio del caso general.

Todavía puede hacerse la siguiente crítica general del esquema completo: al apartarnos del determinismo de la teoría clásica hemos introducido una complicación considerable en la descripción de los fenómenos naturales que no es deseable en absoluto. Esta complicación es innegable, pero queda compensada por la gran simplificación que nos ofrece el *principio general de superposición de los estados* que estudiaremos a continuación. Pero antes es necesario precisar el concepto clave de 'estado' de un sistema atómico general.

Sea un sistema atómico cualquiera, compuesto de partículas o de cuerpos con propiedades dadas (masa, momento de inercia, etc.) que interaccionan de acuerdo con leyes de fuerza dadas. Existirán diversos movimientos posibles de las partículas o de los cuerpos compatibles con las leyes de fuerza. Cada uno de tales movimientos recibe el nombre de *estado* del sistema. Según las ideas clásicas, se podría especificar un estado dando el valor numérico de todas las coordenadas y velocidades de los diversos componentes del sistema en un instante de tiempo, quedando así completamente determinado el movimiento de todo el sistema. Pero los razonamientos de § 3 y § 4 nos indican que no podemos observar un sistema *pequeño* con tanto detalle como supone la teoría clásica. La limitación del poder de observación implica una disminución del número de datos que pueden atribuirse a un sistema. Por tanto, el estado de un sistema atómico tiene que caracterizarse por menos datos o por datos más imprecisos que un conjunto completo de valores numéricos de todas las coordenadas y velocidades en un instante de tiempo particular. Cuando el sistema está constituido por un único fotón, el estado estará completamente especificado dando un estado de traslación en el sentido de § 3 y un estado de polarización en el sentido de § 2.

Puede definirse el estado de un sistema como un movimiento inalterado restringido por tantas condiciones o datos como sea posible teóricamente sin que se interfieran o se contradigan mutuamente. En la práctica, estas condiciones se pueden imponer mediante una preparación adecuada del sistema, que consista, por ejemplo, en hacerle pasar a través de una serie

de aparatos, tales como rendijas y polarímetros, y dejando el sistema inalterado después de la preparación. La palabra 'estado' se emplea indistintamente para designar el estado en un instante particular (después de la preparación) y el estado durante todo el tiempo posterior a la preparación. Para diferenciar ambos significados llamaremos a este último 'estado de movimiento' cuando pudiera dar lugar a confusión.

El principio general de superposición de la mecánica cuántica se aplica a los estados de todo sistema dinámico con cualquiera de los dos significados precedentes. Según dicho principio existe una relación peculiar entre los estados, de forma que cuando un sistema está en un estado bien definido, a la vez se puede considerar que está parcialmente en cada uno de una serie de estados. El estado original debe considerarse como el resultado de una *superposición* de esta serie de estados, lo que es inconcebible desde el punto de vista clásico. Es evidente que existen infinitas maneras de realizar dicha superposición para un mismo estado dado. Recíprocamente, todo conjunto de dos o más estados puede superponerse para dar lugar a un nuevo estado. El procedimiento de expresar un estado como el resultado de una superposición de un conjunto de otros estados es un procedimiento matemático que está siempre permitido, independientemente de toda referencia a las condiciones físicas, como ocurre con el método de descomposición de una onda en componentes de Fourier. El que sea útil en un caso particular dependerá de las condiciones físicas particulares del problema que estemos tratando.

En las dos secciones anteriores se han dado ejemplos del principio de superposición aplicado a un sistema constituido por un único fotón. En § 2 se consideran estados que sólo difieren respecto a la polarización, y en § 3 estados que difieren únicamente respecto al movimiento del fotón como un todo.

La relación entre los estados de cualquier sistema que implica el principio de superposición es de tal naturaleza que no es posible expresarla mediante los conceptos físicos habituales. Desde el punto de vista clásico, no es posible imaginar que un sistema esté parcialmente en cada uno de dos estados dados, y que esto sea equivalente a estar completamente en otro estado distinto. Ello implica una idea completamente nueva a la que nos debemos ir acostumbrando, y sobre cuya base tenemos que construir una teoría matemática exacta sin disponer de ninguna imagen clásica precisa.

Un estado formado por superposición de otros dos tendrá propiedades que, en cierto sentido, serán intermedias entre las de ambos sistemas de partida, aproximándose en mayor o menor grado a las de cada uno, según se le haya atribuido un 'peso' mayor o menor en el proceso de superposición. El nuevo estado estará completamente determinado por los dos estados de partida, cuando se conozcan sus pesos relativos en el proceso de superposición y una diferencia de fase; el significado preciso de pesos y fases en general nos lo proporcionará la teoría matemática. En el caso de la polari-

zación de fotones, sus significados son los que nos dan la óptica clásica, y así por ejemplo, cuando se superponen dos estados polarizados según direcciones perpendiculares con igual peso, el estado resultante puede estar polarizado circularmente en todas direcciones, rectilíneamente según un ángulo $\frac{\pi}{4}$ o elípticamente, según la diferencia de fase.

La naturaleza no clásica del proceso de superposición se pone claramente de manifiesto al considerar dos estados A y B, tales que exista una observación que aplicada al sistema en el estado A dé siempre el resultado particular a y aplicada al sistema en el estado B dé siempre el mismo resultado particular b distinto de a. ¿Cuál será el resultado de la observación aplicada al sistema en el estado resultante de la superposición? La respuesta es que unas veces obtendremos el resultado a y otras el b, según una ley probabilística que depende únicamente de los pesos relativos de A y B en el proceso de superposición. En ningún caso obtendremos un resultado distinto de a o b. *El carácter intermedio del estado formado por la superposición reside en el hecho de que la probabilidad de obtener un resultado particular en una observación es intermedia entre las probabilidades correspondientes a los estados de partida,* y no en que el mismo resultado sea intermedio entre los resultados correspondientes a dichos estados.*

Vemos, pues, que un abandono tan radical de las ideas clásicas, como el afirmar la existencia de relaciones de superposición entre los estados, sólo ha sido posible al tener en cuenta explícitamente la importancia de la alteración que acompaña a toda observación y la incertidumbre consiguiente en el resultado. Cuando se lleva a cabo una observación de un sistema atómico, en general no está determinado el resultado, es decir, que si repetimos el experimento varias veces bajo idénticas condiciones, podemos obtener diversos resultados. Sin embargo, es una ley de la naturaleza el hecho de que si repetimos el experimento un gran número de veces, cada resultado particular aparece en una fracción bien definida del número total de veces, y por tanto, existe una *probabilidad* determinada de obtenerlo. Esta probabilidad es lo que la teoría nos permite calcular. Únicamente está determinado el resultado del experimento en ciertos casos especiales, cuando la probabilidad de obtener un resultado es la unidad.

La hipótesis de que existan relaciones de superposición entre los estados nos lleva a establecer una teoría matemática en la cual las ecuaciones que definen el estado son lineales en las incógnitas. Como consecuencia de ello, se ha intentado establecer analogías con sistemas de la mecánica clásica,

---

* En general, la probabilidad de obtener un resultado particular con el sistema en el estado formado por superposición, no es siempre intermedia entre las probabilidades de los sistemas originales si éstas no son cero o uno; así pues existen limitaciones del 'carácter intermedio' del estado formado por superposición.

tales como cuerdas y membranas vibrantes, cuyas ecuaciones son también lineales y que, por lo tanto, gozan de un principio de superposición. Estas analogías han dado lugar al nombre de 'mecánica ondulatoria' con el que se denomina a veces la mecánica cuántica. Pero es importante recordar que *la superposición que aparece en mecánica cuántica es de una naturaleza esencialmente distinta de las que encontramos en la teoría clásica*, como lo demuestra el hecho de que el principio cuántico de superposición exige que los resultados de observación estén indeterminados a fin de poder dar una interpretación física razonable. Por lo tanto, las analogías pueden prestarse a equívocos.

5. *Formulación matemática del principio de superposición*

Durante el presente siglo ha tenido lugar un profundo cambio en las opiniones que los científicos tenían sobre los principios matemáticos de la física. Anteriormente se suponía que la mecánica de Newton constituía la base para la descripción de todos los fenómenos físicos, y que por tanto, todo físico teórico debía contribuir al desarrollo y aplicación de estos principios. El reconocimiento de que no existe razón lógica alguna para que los principios newtonianos y otros principios clásicos tengan que ser válidos fuera del campo en que han sido comprobados experimentalmente, trajo consigo las primeras experiencias que pusieron de manifiesto la absoluta necesidad de apartarse de estos principios. Los intentos de superar las divergencias se concretaron en la introducción de nuevos formalismos matemáticos, de nuevos sistemas de axiomas y reglas en los métodos de la física teórica.

La mecánica cuántica es un buen ejemplo de las nuevas ideas. En ella se exige que los estados de un sistema dinámico estén relacionados con las variables dinámicas de una forma nueva, ininteligible desde el punto de vista clásico. Los estados y las variables dinámicas deben estar representadas por cantidades matemáticas de naturaleza diferente a las que se utilizan corrientemente en física. El nuevo esquema constituirá una teoría física precisa cuando se hayan especificado todos los axiomas y reglas que rigen las cantidades matemáticas se hayan dado ciertas leyes que relacionen los hechos físicos con dichas cantidades, de tal manera que de unas condiciones físicas dadas puedan inferirse ecuaciones matemáticas y recíprocamente. En la aplicación de la teoría deberíamos disponer de cierta información física, que procederíamos a expresar en forma de ecuaciones entre las cantidades matemáticas. Con ayuda de los axiomas y las reglas de manipulación deduciríamos nuevas ecuaciones y acabaríamos interpretándolas como condiciones físicas. La justificación de todo el esquema depende, aparte de su coherencia intrínseca, de la concordancia de los resultados finales con los experimentos.

~~mecánica~~ la mecánica ondulatoria no es pues que un caso ~~muy~~ particular de una formulación más general de la mecánica cuántica. La mecánica ondulatoria sólo ~~llenaría~~ representación de Schrödinger. (pag 99)

Empezaremos a construir el esquema considerando las relaciones matemáticas entre los estados de un sistema dinámico en un instante de tiempo que se derivan de la formulación matemática del principio de superposición. La superposición es un cierto proceso aditivo, e implica que los estados puedan sumarse de algún modo para dar nuevos estados. Por lo tanto, los estados tienen que estar asociados con cantidades matemáticas que puedan sumarse entre sí para dar cantidades de la misma clase. Las cantidades matemáticas más sencillas que disfrutan de esta propiedad son los vectores. Los vectores ordinarios definidos en un espacio de un número finito de dimensiones no son suficientemente generales para la mayoría de los sistemas dinámicos de la mecánica cuántica. Nos vemos obligados a generalizar los vectores a un espacio de infinitas dimensiones, con lo que el tratamiento matemático se hace complicado por razones de convergencia. Sin embargo, de momento únicamente vamos a considerar propiedades que puedan ser deducidas sobre la base de un conjunto sencillo de axiomas, y dejaremos a un lado las cuestiones de convergencia y los temas relacionados con ella hasta que nos veamos obligados a tenerlos en cuenta.

Es conveniente designar con un nombre especial a los vectores que se asocian a los estados de un sistema en mecánica cuántica, tanto si forman parte de un espacio de un número finito de dimensiones como de un espacio de infinitas dimensiones. Las denominaremos *vectores ket*, o simplemente *kets*, y los representaremos con el símbolo especial $|\rangle$. Si queremos especificar un ket particular mediante una letra, por ejemplo la A, la colocaremos entre los dos signos así $|A\rangle$. La conveniencia de esta notación aparecerá clara cuando hayamos desarrollado todo el esquema.

Los vectores ket se pueden multiplicar por números complejos y pueden también sumarse entre ellos para dar nuevos kets; por ejemplo, de los dos kets $|A\rangle$ y $|B\rangle$ podemos formar

$$c_1 |A\rangle + c_2 |B\rangle = |R\rangle, \tag{1}$$

en donde c_1 y c_2 son números complejos arbitrarios. También es posible efectuar con ellos operaciones lineales más generales, como sumar una serie infinita de kets, o si tenemos un ket $|x\rangle$ que dependa de un parámetro x que puede tomar todos los valores de un cierto intervalo, integrar respecto a x para obtener un nuevo ket

$$\int |x\rangle \, dx = |Q\rangle$$

Todo vector ket que se pueda expresar linealmente en función de los otros se dice que es *dependiente* de ellos. Se dice que un conjunto de kets es *independiente*, si ninguno de ellos se puede expresar linealmente en función de los otros.

Ahora hacemos la hipótesis de que *a cada estado de un sistema dinámico en un instante particular le corresponde un ket, siendo la correspondencia tal que si un estado está definido como superposición de otros dos, su*

correspondiente ket puede expresarse linealmente en función de los kets correspondientes a dichos estados y recíprocamente. Por tanto, el estado R resulta de una superposición de los estados A y B si los correspondientes kets están ligados por (1).

La hipótesis anterior nos lleva a introducir ciertas propiedades del proceso de superposición, que de hecho son necesarias para que sea apropiada la palabra 'superposición'. Cuando se superponen dos o más estados, no importa el orden en que entran en el proceso de superposición, y así dicho proceso es simétrico respecto a los estados que se superponen. Además, de (1) deducimos que (exceptuando el caso en que c_1 o c_2 sean nulos) si el estado R puede formarse por superposición de los estados A y B, el estado A puede formarse por superposición de B y R, y el B por superposición de A y R. Las relaciones de superposición son, pues, simétricas respecto a los tres estados A, B y R.

Cuando un estado está formado por la superposición de otros dos se dice que es *dependiente* de ellos. Más en general, se dice que un estado es *dependiente* de un conjunto finito o infinito de otros estados, si su correspondiente ket es dependiente de los kets que corresponden a dichos estados. Se dice que un conjunto de estados es *independiente* si ninguno de los estados es dependiente de los otros.

Para proseguir con la formulación matemática del principio de superposición tenemos que introducir una nueva hipótesis, y afirmar que por superposición de un estado consigo mismo no podemos construir ningún estado nuevo, sino que siempre obtenemos el mismo estado. Si el estado original correspondía al ket $|A\rangle$, al superponerle consigo mismo el estado resultante corresponde a

$$c_1 |A\rangle + c_2 |A\rangle = (c_1 + c_2) |A\rangle ,$$

donde c_1 y c_2 son números. Puede ocurrir que $c_1 + c_2 = 0$, en cuyo caso el resultado de la superposición no representa nada en absoluto, pues las dos componentes se han eliminado mutuamente por un efecto de interferencia. Nuestra nueva hipótesis exige que, salvo en este caso particular, el estado resultante sea el mismo que el original, y por tanto, $(c_1 + c_2)|A\rangle$ tiene que corresponder al mismo estado al que corresponde $|A\rangle$. Pero $c_1 + c_2$ es un número complejo arbitrario, de donde deducimos que *si multiplicamos el ket que corresponde a un estado por un número complejo arbitrario distinto de cero, el ket resultante corresponderá al mismo estado.* Así, pues, un estado viene caracterizado por la dirección de un ket, independientemente de la longitud que le atribuyamos. Los estados de un sistema dinámico están en correspondencia biyectiva con las posibles direcciones de los vectores ket, considerando como una sola las direcciones de $|A\rangle$ y de $-|A\rangle$.

Esta hipótesis nos muestra claramente la diferencia fundamental entre la superposición que se da en mecánica cuántica y cualquier otra super-

posición clásica. En un sistema clásico para el que sea válido un principio de superposición, como por ejemplo una membrana vibrante, cuando se superpone un estado consigo mismo el resultado es un estado *diferente*, cuya amplitud de oscilación es distinta. No existe ninguna característica física en los estados de un sistema cuántico que corresponda a la amplitud de las oscilaciones clásicas, más que la relación entre las amplitudes en los distintos puntos de la membrana. Además, si bien existe un estado clásico de amplitud nula en todos los puntos de la membrana, que es el estado de reposo, no hay ninguno correspondiente a éste en un sistema cuántico, pues el ket nulo no corresponde a ningún estado de existencia.

Dados dos estados que correspondan a dos kets $|A\rangle$ y $|B\rangle$, el estado más general que se puede formar por superposición de ambos corresponde a un ket $|R\rangle$ que viene determinado dando dos números complejos: los coeficientes c_1 y c_2 de la ecuación (1). Si multiplicamos los dos coeficientes por un mismo factor (también complejo), el ket $|R\rangle$ quedará multiplicado por dicho factor y el estado correspondiente será el mismo de antes. Por lo tanto, únicamente interviene en la determinación del estado R el cociente entre ambos coeficientes. Así pues, dicho estado queda determinado por un número complejo, o lo que es lo mismo, por dos parámetros reales. Por lo tanto, dados dos estados, por superposición de ambos podemos formar una doble infinidad de estados.

Este resultado viene confirmado por los ejemplos explicados en § 2 y § 3. En el ejemplo de § 2 existen dos únicos estados de polarización de un fotón independientes, pudiéndose elegir como tales dos estados de polarización rectilínea según direcciones paralela y perpendicular a una dada; por superposición de ambos se puede formar una doble infinidad de estados de polarización, a saber, todos los estados de polarización elíptica, para cuya especificación en el caso general son necesarios dos parámetros. Asimismo, en el ejemplo del § 3, por superposición de dos estados de traslación de un fotón dados, podemos obtener una doble infinidad de nuevos estados de traslación, pues el estado general así formado depende de dos parámetros, que pueden ser, por ejemplo, la relación de amplitudes de las dos funciones de onda que deben sumarse y su diferencia de fase relativa. Esta confirmación muestra la necesidad de tomar coeficientes complejos en la ecuación (1). Si únicamente admitiéramos coeficientes reales, puesto que una vez conocidos $|A\rangle$ y $|B\rangle$ únicamente interviene el cociente de ambos coeficientes para especificar la dirección del ket $|R\rangle$ resultante, solamente podríamos construir una simple infinidad de estados por superposición de $|A\rangle$ y $|B\rangle$.

6. *Vectores bra y ket*

Siempre que tenemos un conjunto de vectores en una teoría matemática cualquiera podemos construir un segundo conjunto de vectores, que los

matemáticos denominan vectores duales. Vamos a describir el método de obtenerlo en el caso de que los vectores de partida sean nuestros kets.

Sea un número ϕ función de un ket $|A\rangle$; es decir, que a cada ket $|A\rangle$ le corresponde un número ϕ, y además exijamos que la función sea lineal, lo que quiere decir que el número que corresponde a $|A\rangle + |A'\rangle$ es igual a la suma de los números que corresponden a $|A\rangle$ y a $|A'\rangle$, y que el número que corresponde a $c|A\rangle$ es igual a c multiplicado por el número que corresponde a $|A\rangle$, siendo c un factor numérico arbitrario. Entonces, el número ϕ que corresponde a $|A\rangle$ puede ser considerado como el producto escalar de dicho $|A\rangle$ con un nuevo vector, existiendo tantos vectores nuevos como funciones lineales de los vectores ket. La justificación de que podamos considerar a ϕ de este modo, reside, como veremos más adelante (véanse las ecuaciones (5) y (6)), en el hecho de que los nuevos vectores pueden sumarse entre ellos y multiplicarse por números dando nuevos vectores de la misma clase. A pesar de que los nuevos vectores sólo están dados cuando conocemos los productos escalares con los vectores de partida, esto nos basta para construir una teoría matemática con ellos.

Les denominaremos vectores bra o simplemente bras, y los representaremos con el símbolo $\langle|$, imagen simétrica del símbolo de un ket. Cuando queramos especificar un bra particular mediante una letra B, la escribiremos entre los dos signos así $\langle B|$. El producto escalar del bra $\langle B|$ y del ket $|A\rangle$ lo escribiremos $\langle B|A\rangle$, es decir, yuxtaponiendo los símbolos del bra y del ket, con el bra a la izquierda, y contrayendo las dos líneas verticales en una sola para abreviar.

Podemos considerar los símbolos $\langle$ y $\rangle$ como tipos característicos de paréntesis. Un producto escalar $\langle B|A\rangle$ aparece ahora como una expresión entre paréntesis completos, * mientras que un bra $\langle B|$ o un ket $|A\rangle$ son expresiones entre paréntesis incompletos. Por lo tanto, son válidas las siguientes reglas: *toda expresión entre paréntesis completos expresa un número, y toda expresión entre paréntesis incompletos expresa un vector, que será bra o ket según tenga el primer paréntesis o el segundo.*

La condición de que el producto escalar de $\langle B|$ y $|A\rangle$ sea una función lineal de $|A\rangle$ se puede escribir simbólicamente así:

$$\langle B|\{|A\rangle + |A'\rangle\} = \langle B|A\rangle + \langle B|A'\rangle\,, \tag{2}$$

$$\langle B|\{c|A\rangle\} = c\langle B|A\rangle\,, \tag{3}$$

donde c es un número.

Se considera que un bra está completamente definido, cuando se conoce su producto escalar con cualquier ket, de modo que si el producto escalar

---

* Los nombres de bra y ket, introducidos por Dirac, cuando se yuxtaponen forman la palabra 'bracket' que en inglés significa paréntesis. *(N. del T.)*

de un determinado bra con todo ket es nulo, debe considerarse también nulo el bra. En símbolos, si

$$\left.\begin{array}{ll} & \langle P|A\rangle = 0, \quad \text{para todo } |A\rangle, \\ \text{entonces} & \langle P| = 0. \end{array}\right\} \quad (4)$$

Definimos la suma de dos bras $\langle B|$ y $\langle B'|$ por la condición de que su producto escalar con cualquier ket $|A\rangle$ sea igual a la suma de los productos escalares de $\langle B|$ y $\langle B'|$ con $|A\rangle$,

$$\{\langle B| + \langle B'|\}|A\rangle = \langle B|A\rangle + \langle B'|A\rangle, \quad (5)$$

y el producto de un bra $\langle B|$ con un número c por la condición de que su producto escalar con cualquier ket $|A\rangle$ sea igual a c multiplicado por el producto escalar de $\langle B|$ y $|A\rangle$,

$$\{c\langle B|\}|A\rangle = c\langle B|A\rangle. \quad (6)$$

Las ecuaciones (2) y (5) expresan que el producto escalar de un bra y un ket verifica la propiedad distributiva de la multiplicación, y las (3) y (6) indican que la multiplicación por factores numéricos verifica las propiedades algebraicas ordinarias.

Los vectores bra, tal como los hemos introducido aquí, son de distinta naturaleza que los vectores ket, y hasta aquí, aparte del producto escalar, no existe ninguna otra relación entre bras y kets. Ahora hacemos la hipótesis de que *existe una correspondencia biyectiva entre los bras y los kets, tal que el bra que corresponde a* $|A\rangle + |A'\rangle$ *es igual a la suma de los bras correspondientes a* $|A\rangle$ *y* $|A'\rangle$ *y el bra correspondiente a* $c|A\rangle$ *es igual a* $\bar{c}$ *multiplicado por el bra correspondiente a* $|A\rangle$, *siendo* $\bar{c}$ *el complejo conjugado de* c. Emplearemos la misma letra para indicar un ket y su bra correspondiente. Así el bra que corresponde a $|A\rangle$ es $\langle A|$.

La relación que existe entre un ket y su correspondiente bra justifica el que llamemos a cada uno el conjugado imaginario del otro. Nuestros bras y kets son cantidades complejas, puesto que se pueden multiplicar por números complejos obteniéndose vectores de la misma naturaleza que antes; pero son cantidades complejas especiales, ya que no pueden separarse en parte real y parte imaginaria pura. El procedimiento habitual de obtener la parte real de una cantidad compleja, que consiste en tomar la semisuma de dicha cantidad y su conjugada, es inaplicable debido a que los bras y los kets son de distinta naturaleza y no se pueden sumar unos con otros. Para llamar la atención sobre esta diferencia, utilizaremos las palabras 'complejo conjugado' cuando nos refiramos a números u otras cantidades complejas que puedan separarse en parte real y parte imaginaria pura, y las de 'imaginario conjugado' para las que no disfruten de esta propiedad. Para las cantidades del primer tipo, cuando queramos indicar el complejo conjugado de una cantidad dada, emplearemos la notación que consiste en colocar una raya sobre dicha cantidad.

Teniendo en cuenta la correspondencia biyectiva entre bras y kets, *todo estado de nuestro sistema dinámico en un instante particular puede especificarse tanto por la dirección de un bra como por la de un ket.* De hecho toda la teoría será simétrica en lo esencial respecto a bras y kets.

Dados dos kets $|A\rangle$ y $|B\rangle$, podemos formar con ellos un número $\langle B|A\rangle$, producto escalar del primero con el imaginario conjugado del segundo. Tal número depende linealmente de $|A\rangle$ y antilinealmente de $|B\rangle$. Dependencia antilineal significa que el número que se obtiene con $|B\rangle + |B'\rangle$ es igual a la suma de los números que se obtienen con $|B\rangle$ y con $|B'\rangle$, y que el número que se obtiene con $c|B\rangle$ es igual a $\bar{c}$ multiplicado por el número que se obtiene con $|B\rangle$. Hay otro procedimiento de formar un número que dependa linealmente de $|A\rangle$ y antilinealmente de $|B\rangle$, que consiste en tomar el número complejo conjugado del producto escalar de $|B\rangle$, con el conjugado imaginario de $|A\rangle$. *Supondremos que estos dos números son iguales,* es decir

$$\langle B|A\rangle = \overline{\langle A|B\rangle} \tag{7}$$

Poniendo aquí $|B\rangle = |A\rangle$, resulta que el número $\langle A|A\rangle$ es real. Hacemos además la nueva hipótesis de que

$$\langle A|A\rangle > 0, \tag{8}$$

salvo si $|A\rangle = 0$.

En el espacio ordinario, dados dos vectores, se puede formar con ellos un número — su producto escalar — que es real y simétrico respecto a ellos. En el espacio de los vectores bra o en el de los vectores ket, dados dos vectores cualesquiera también podemos formar un número — el producto escalar de uno de ellos con el imaginario conjugado del otro — que ahora es complejo, y se transforma en su complejo conjugado al cambiar el orden de los vectores. Existe, por tanto, un tipo de perpendicularidad en estos espacios, que es una generalización de la perpendicularidad en el espacio ordinario. Diremos que un bra y un ket son *ortogonales* si su producto escalar es nulo, y que dos bras o dos kets son ortogonales, si el producto escalar de uno de ellos por el imaginario conjugado del otro es cero. Además, diremos que dos estados de nuestro sistema dinámico son ortogonales, si sus vectores correspondientes también lo son.

Definimos la *longitud* de un bra $\langle A|$ o de su ket imaginario conjugado $|A\rangle$ como la raíz cuadrada del número positivo $\langle A|A\rangle$. Si dado un estado queremos caracterizarlo mediante un bra o un ket, únicamente queda determinada la dirección del vector, pero el vector queda indeterminado en un factor numérico arbitrario. A menudo es conveniente elegir dicho factor de forma que el vector tenga longitud unidad. Este procedimiento se llama *normalización* y el vector así elegido se dice que está *normalizado.* Sin embargo, aún no está completamente determinado el vector, pues se le puede multiplicar por cualquier factor de módulo uno, es decir, cualquier

número de la forma $e^{i\gamma}$ con γ real, sin modificar su longitud. A este número le llamaremos *factor de fase*.

Las hipótesis precedentes nos dan el esquema completo de las relaciones entre los estados de un sistema dinámico en un instante particular. Estas relaciones aparecen en forma matemática, pero implican condiciones físicas que nos llevarán, cuando hayamos desarrollado más la teoría, a resultados que se pueden expresar en términos de los datos de observación. Por ejemplo, el que dos estados sean ortogonales, no implica de momento más que una determinada ecuación en nuestro formalismo, pero tal ecuación implica a su vez una relación física definida entre los estados, que los próximos desarrollos de la teoría nos permitirán interpretar en función de los datos de observación (véase el final de la pág. 36 y principio de la 37).

$\psi_n(\xi) = \langle \xi | n \rangle$

II

VARIABLES DINÁMICAS Y OBSERVABLES

7. *Operadores lineales*

En la sección anterior hemos considerado funciones lineales numéricas de kets, que nos han llevado al concepto de bra. Ahora consideraremos funciones lineales vectoriales de kets, que nos llevarán al concepto de operador lineal.

Sea un ket $|F\rangle$ función de un ket $|A\rangle$, es decir, que a cada ket $|A\rangle$ le corresponde un ket $|F\rangle$, y supongamos además que dicha función sea lineal, o lo que es lo mismo, que el ket $|F\rangle$ que corresponde a $|A\rangle + |A'\rangle$ es igual a la suma de los kets $|F\rangle$ correspondientes a $|A\rangle$ y a $|A'\rangle$, y que el ket $|F\rangle$ que corresponde a $c|A\rangle$ es igual a c multiplicado por el ket $|F\rangle$ correspondiente a $|A\rangle$, donde c es un factor numérico arbitrario. Bajo estas condiciones, podemos considerar el paso de $|A\rangle$ a $|F\rangle$ como la aplicación de un *operador lineal* a $|A\rangle$. Si introducimos el símbolo $\hat{\alpha}$ para el operador lineal, se puede escribir

$$|F\rangle = \hat{\alpha}|A\rangle,$$

donde indicamos la aplicación de α sobre $|A\rangle$ como un producto de α por $|A\rangle$. Establecemos la regla de que en tales productos *el ket debe estar siempre a la derecha del operador lineal.* Podemos expresar ahora las condiciones de linealidad antes mencionadas mediante las ecuaciones

$$\left.\begin{aligned} \alpha\{|A\rangle + |A'\rangle\} &= \alpha|A\rangle + \alpha|A'\rangle, \\ \alpha\{c|A\rangle\} &= c\alpha|A\rangle. \end{aligned}\right\} \quad (1)$$

Diremos que un operador lineal está completamente definido, si conocemos el resultado de su aplicación sobre cada ket. Por tanto, debemos considerar nulo a un operador lineal que aplicado a cualquier ket da cero, y asimismo diremos que dos operadores lineales son iguales, si dan el mismo resultado al aplicarlos a cualquier ket.

También podemos sumar operadores, y definimos la suma de dos de ellos como aquel operador lineal que aplicado a cualquier ket da un resul-

tado igual a la suma de los kets que se obtienen al aplicar cada uno de los operadores por separado a dicho ket. Así pues definimos $\alpha + \beta$ por

$$\{\alpha + \beta\}|A\rangle = \alpha|A\rangle + \beta|A\rangle \tag{2}$$

para todo $|A\rangle$. La ecuación (2) y la primera de las ecuaciones (1) nos indican que el producto de operadores lineales por kets verifica la propiedad distributiva de la multiplicación.

Asimismo podemos multiplicar operadores lineales entre sí, y definimos el producto de dos operadores como aquel operador lineal que aplicado a cualquier ket da el mismo resultado que se obtiene al aplicar sucesivamente los dos operadores dados a dicho ket. Así pues, hemos definido el producto $\alpha\beta$ como el operador lineal que aplicado a cualquier ket $|A\rangle$, lo transforma en el ket que se obtiene aplicando primero β a $|A\rangle$, y después α al resultado. En símbolos

$$\{\alpha\beta\}|A\rangle = \alpha\{\beta|A\rangle\}.$$

Esta definición es equivalente a la propiedad asociativa de la multiplicación del triple producto de α, β y $|A\rangle$, y nos permite por lo tanto escribir dicho producto sin paréntesis $\alpha\beta|A\rangle$. Sin embargo, este triple producto no da en general el mismo resultado que se obtendría aplicando primero α a $|A\rangle$ y después β al resultado, es decir, que en general $\alpha\beta|A\rangle$ es distinto de $\beta\alpha|A\rangle$, y por tanto, en general, $\alpha\beta$ es distinto de $\beta\alpha$. *La propiedad conmutativa no es válida para la multiplicación de operadores lineales.* Puede ocurrir que en casos especiales dos operadores lineales ξ y η sean tales que $\xi\eta$ y $\eta\xi$ sean iguales. En tal caso diremos que ξ *conmuta con* η, o que ξ y η *conmutan.*

Combinando las operaciones dadas de suma y multiplicación de operadores lineales se pueden formar sumas y productos con más de dos operadores, y así podemos construir con ellos un álgebra. En dicho álgebra no será válida la propiedad conmutativa de la multiplicación, y podrá ocurrir también que el producto de dos operadores lineales sea nulo sin que lo sea ninguno de los dos factores. Pero todas las demás propiedades del álgebra ordinaria serán válidas, incluidas las propiedades asociativa y distributiva de la multiplicación, como puede comprobarse fácilmente.

Si tomamos un número k y lo multiplicamos por vectores ket, aparece como un operador lineal que se aplica a kets, pues verifica las condiciones (1) substituyendo k por α. Por tanto, un número es un caso particular de operador lineal. Tiene la propiedad de que conmuta con todo otro operador lineal, lo que le distingue de un operador lineal general.

Hasta aquí únicamente hemos aplicado los operadores lineales a kets. Pero también pueden ser aplicados a bras, con el significado siguiente. Tomemos el producto escalar del bra $\langle B|$ con el ket $\alpha|A\rangle$. Tal producto escalar es un número que depende linealmente de $|A\rangle$ y por lo tanto, dada la definición de bra, puede ser considerado como el producto escalar de $|A\rangle$ por un cierto bra. Dicho bra depende linealmente de $\langle B|$, así que

puede considerarse como el resultado de aplicar cierto operador lineal a $\langle B|$. Dicho operador lineal está determinado unívocamente por el operador lineal α de antes y es lógico decir que se trata del mismo operador lineal que actúa ahora sobre un vector bra.

Una notación conveniente para indicar el bra que resulta de aplicar α al bra $\langle B|$ es $\langle B|\alpha$, y en dicha notación, la ecuación que define $\langle B|\alpha$ es

$$\{\langle B|\alpha\}|A\rangle = \langle B|\{\alpha|A\rangle\} \tag{3}$$

para cualquier $|A\rangle$, que expresa simplemente la propiedad asociativa de la multiplicación para el triple producto de $\langle B|$, α y $|A\rangle$. Establecemos la regla general de que en todo producto de un bra y un operador lineal, siempre ha de estar el bra a la izquierda. Ahora ya podemos escribir el triple producto de $\langle B|$, α y $|A\rangle$ simplemente $\langle B|\alpha|A\rangle$ sin ningún paréntesis. Puede comprobarse fácilmente que la propiedad distributiva de la multiplicación es válida para los productos de operadores lineales y kets.

En nuestro esquema cabe todavía un nuevo tipo de producto, que es el producto de un ket y un bra con el ket a la izquierda $|A\rangle\langle B|$. Para ver qué significa dicho producto, multipliquémoslo por un ket $|P\rangle$ arbitrario, que colocaremos a la derecha, y supongamos que es válida la propiedad asociativa para esta multiplicación. Dicho producto vale entonces $|A\rangle\langle B|P\rangle$, o sea el ket $|A\rangle$ multiplicado por el número $\langle B|P\rangle$, y es función lineal del ket $|P\rangle$. Por tanto, $|A\rangle\langle B|$ aparece como un operador lineal que puede actuar sobre kets. También puede actuar sobre bras, siendo su producto por un bra $\langle Q|$ a la izquierda $\langle Q|A\rangle\langle B|$, que es igual al número $\langle Q|A\rangle$ multiplicado por el bra $\langle B|$. Debemos distinguir claramente entre los productos $|A\rangle\langle B|$ y $\langle B|A\rangle$ de los mismos factores pero en orden inverso, siendo el último producto un número, como ya sabemos.

Disponemos ahora de un esquema algebraico completo referente a tres clases de cantidades: vectores bra, vectores ket y operadores lineales. Podemos multiplicarlos entre sí de todas las formas indicadas, siendo válidas las propiedades asociativa y distributiva de la multiplicación en todos los casos, pero no la propiedad conmutativa. En el esquema general sigue siendo válida la regla de notación dada en la sección anterior, de que toda expresión entre paréntesis completos, que tenga $\langle$ a la izquierda y $\rangle$ a la derecha, representa un número, mientras que toda expresión entre paréntesis incompletos, que no tenga más que $\langle$ o $\rangle$, representa un vector.

Respecto al significado físico del esquema, hemos supuesto siempre que los bras y los kets, o mejor dicho las direcciones de dichos vectores, corresponden a los estados de un sistema dinámico en un instante particular. Ahora hacemos la nueva hipótesis de que *los operadores lineales corresponden a las variables dinámicas en dicho instante*. Bajo el nombre de variables dinámicas entendemos cantidades como las coordenadas, las componentes de la velocidad, del momento y del momento angular de las partículas así como funciones de éstas — de hecho las variables que se emplean

en mecánica clásica —. La nueva hipótesis exige que dichas cantidades aparezcan también en mecánica cuántica, pero con la notable diferencia de que en ella *están regidas por un álgebra en la cual no es válida la propiedad conmutativa de la multiplicación.*

El hecho de que el álgebra de las variables dinámicas sea distinta en ambas teorías es una de las diferencias más importantes entre la mecánica cuántica y la clásica. Más adelante veremos, a este respecto, que a pesar de esta diferencia fundamental, las variables dinámicas de la mecánica cuántica siguen gozando de muchas propiedades comunes con sus equivalentes clásicas, y que es posible construir una teoría análoga en gran parte a la clásica que constituye una excelente generalización de ella.

Es conveniente utilizar la misma letra para especificar una variable dinámica y su correspondiente operador lineal. De hecho, consideraremos a ambos como una sola cosa, sin que ello dé lugar a confusión.

8. *Relaciones conjugadas*

Nuestros operadores lineales son cantidades complejas, puesto que se pueden multiplicar por números complejos resultando nuevas cantidades de la misma naturaleza. Así pues han de corresponder en general a variables dinámicas complejas, es decir, a funciones complejas de las coordenadas, velocidades, etc. A fin de poder ver qué clase de operador lineal corresponde a una variable dinámica real, precisamos una elaboración más amplia de la teoría.

Consideremos el ket conjugado imaginario de $\langle P|\alpha$. Dicho ket depende antilinealmente de $\langle P|$ y, por tanto, depende linealmente de $|P\rangle$. Por consiguiente, puede ser considerado como el resultado de aplicar un cierto operador lineal a $|P\rangle$. A este operador lineal se le llama *adjunto* de α y le designaremos por $\bar{\alpha}$. Con esta notación, el conjugado imaginario de $\langle P|\alpha$ es $\bar{\alpha}|P\rangle$.

En la fórmula (7) del capítulo I pongamos $\langle P|\alpha$ en lugar de $\langle A|$, y $\bar{\alpha}|P\rangle$ en lugar de $|A\rangle$. El resultado es

$$\langle B|\bar{\alpha}|P\rangle = \overline{\langle P|\alpha|B\rangle} \tag{4}$$

Esta es una fórmula general, válida para todo par de kets $|B\rangle$ y $|P\rangle$ y para todo operador lineal α, que nos expresa una de las propiedades más utilizadas del adjunto.

Substituyendo α por $\bar{\alpha}$ en (4), resulta

$$\langle B|\bar{\bar{\alpha}}|P\rangle = \overline{\langle P|\bar{\alpha}|B\rangle} = \langle B|\alpha|P\rangle,$$

después de haber utilizado otra vez (4) intercambiando $|P\rangle$ con $|B\rangle$. Esto

es válido para todo ket $|P\rangle$, de donde deducimos, con ayuda de (4) del capítulo I,

$$\langle B|\bar{\alpha} = \langle B|\alpha,$$

y puesto que esta ecuación es válida para todo bra $\langle B|$, podemos deducir que

$$\bar{\bar{\alpha}} = \alpha.$$

Luego, *el adjunto del adjunto de un operador lineal es igual al operador lineal de partida.* Esta relación entre un operador y su adjunto es análoga a la que existe entre un número y su complejo conjugado, y además podemos comprobar fácilmente que en el caso particular de que el operador lineal sea un número, su operador lineal adjunto es el número complejo conjugado del dado. Por lo tanto, queda justificado suponer que *el adjunto de un operador lineal corresponde al complejo conjugado de la variable dinámica.* Con este significado físico del adjunto de un operador lineal, podemos llamar al operador adjunto igualmente *operador lineal complejo conjugado,* lo que está de acuerdo con nuestra notación $\bar{\alpha}$.

Un operador lineal puede ser igual a su adjunto, en cuyo caso se dice que es *autoadjunto.* Corresponde a una variable dinámica real, y por esto se le denomina también *operador lineal real.* Todo operador lineal puede ser separado en parte real y parte imaginaria pura. Por esta razón para los operadores lineales se aplica la expresión 'complejo conjugado' y no la de 'imaginario conjugado'.

Evidentemente, el complejo conjugado de la suma de dos operadores lineales es igual a la suma de sus complejos conjugados. Para obtener el complejo conjugado del producto de dos operadores lineales α y β aplicamos la fórmula (7) del capítulo I tomando

$$\langle A| = \langle P|\alpha, \qquad \langle B| = \langle Q|\bar{\beta},$$

de forma que

$$|A\rangle = \bar{\alpha}|P\rangle, \qquad |B\rangle = \beta|Q\rangle.$$

El resultado es

$$\langle Q|\beta\bar{\alpha}|P\rangle = \overline{\langle P|\alpha\beta|Q\rangle} = \langle Q|\overline{\alpha\beta}|P\rangle$$

deducido de (4). Puesto que la fórmula es válida para todo $|P\rangle$ y $\langle Q|$, resulta

$$\beta\bar{\alpha} = \overline{\alpha\beta} \tag{5}$$

Por lo tanto, *el complejo conjugado del producto de dos operadores lineales es igual al producto de los complejos conjugados de los factores en orden inverso.*

Como simple ilustración de este resultado, debe hacerse notar que si ξ y η son reales, en general $\xi\eta$ no es real. Esto representa una diferencia importante con la mecánica clásica. Sin embargo, $\xi\eta + \eta\xi$ sí que es real, y lo mismo ocurre con $i(\xi\eta - \eta\xi)$. Únicamente si ξ y η conmutan, es también

real $\xi\eta$. Otra consecuencia es que si ξ es real, también lo es ξ^2 y más en general ξ^n, siendo n cualquier entero positivo.

Podemos calcular el complejo conjugado del producto de tres operadores lineales aplicando sucesivamente la regla (5) para el complejo conjugado de dos de ellos. Tenemos

$$\overline{\alpha\beta\gamma} = \overline{\alpha(\beta\gamma)} = \overline{\beta\gamma}\,\bar{\alpha} = \bar{\gamma}\,\bar{\beta}\,\bar{\alpha}, \tag{6}$$

Así pues, el complejo conjugado del producto de tres operadores es igual al producto de los complejos conjugados en orden inverso. Esta regla puede hacerse extensiva al producto de un número cualquiera de operadores.

En la sección anterior vimos que el producto $|A\rangle\langle B|$ es un operador lineal. Podemos obtener su complejo conjugado directamente de la definición de adjunto. Multiplicando $|A\rangle\langle B|$ por un bra cualquiera $\langle P|$ obtenemos $\langle P|A\rangle\langle B|$, cuyo imaginario conjugado es

$$\overline{\langle P|A\rangle}|B\rangle = \langle A|P\rangle|B\rangle = |B\rangle\langle A|P\rangle.$$

Por tanto,

$$\overline{|A\rangle\langle B|} = |B\rangle\langle A|. \tag{7}$$

Tenemos ahora distintas leyes que hacen referencia a complejos conjugados y a imaginarios conjugados de productos, a saber, la ecuación (7) del capítulo I, las ecuaciones (4), (5), (6) y (7) del presente capítulo, y la regla de que el imaginario conjugado de $\langle P|\alpha$ es $\bar{\alpha}|P\rangle$. Todas ellas pueden resumirse en una única regla fácil de recodar: *el complejo conjugado o el imaginario conjugado de cualquier producto de vectores bra, vectores ket y operadores lineales se obtiene tomando el complejo conjugado o el imaginario conjugado de cada factor e invirtiendo el orden de todos los factores.* Se puede comprobar fácilmente que esta regla es general, y que es válida también en los casos que no hemos considerado explícitamente.

Teorema. *Si ξ es un operador lineal real, y*

$$\xi^m|P\rangle = 0 \tag{8}$$

para un cierto ket $|P\rangle$, siendo m un entero positivo, entonces

$$\xi|P\rangle = 0.$$

Para demostrar el teorema tomemos en primer lugar el caso de $m = 2$. de la ecuación (8) resulta

$$\langle P|\xi^2|P\rangle = 0,$$

lo que nos dice que el ket $\xi|P\rangle$ multiplicado por su bra imaginario conjugado $\langle P|\xi$ es cero. Según la hipótesis (8) del capítulo I, poniendo $\xi|P\rangle$ en lugar de $|A\rangle$, resulta que $\xi|P\rangle$ tiene que ser nulo. Queda así probado el teorema para $m = 2$.

Ahora consideremos el caso $m > 2$ y llamemos

$$\xi^{m-2}|P\rangle = |Q\rangle.$$

Con esta notación, la ecuación (8) da

$$\xi^2|Q\rangle = 0.$$

Aplicando el teorema para $m = 2$, obtenemos

$$\xi|Q\rangle = 0$$

o su equivalente

$$\xi^{m-1}|P\rangle = 0. \tag{9}$$

Reiterando el procedimiento mediante el que hemos obtenido la ecuación (9) a partir de la (8), obtendremos sucesivamente

$$\xi^{m-2}|P\rangle = 0, \quad \xi^{m-3}|P\rangle = 0, \quad \ldots, \quad \xi^2|P\rangle = 0, \quad \xi|P\rangle = 0,$$

quedando así demostrado el teorema en el caso general.

9. *Autovalores y autovectores*

Es preciso hacer un desarrollo más amplio de la teoría de operadores lineales y estudiar la ecuación

$$\alpha|P\rangle = a|P\rangle, \tag{10}$$

en la que α es un operador lineal y a es un número. Esta ecuación se presenta con frecuencia de forma que α es un operador lineal conocido, y el número a y el ket $|P\rangle$ son las incógnitas, que debemos intentar determinar de modo que satisfagan a la ecuación (10), sin considerar la solución trivial $|P\rangle = 0$.

La ecuación (10) significa que al aplicar el operador α al ket $|P\rangle$, éste queda multiplicado por un factor numérico y no cambia de dirección, o bien que queda multiplicado por el factor cero y deja de tener dirección. Por supuesto, dicho operador α aplicado a otros kets cambiará tanto sus longitudes como sus direcciones. Tengamos presente que en la ecuación (10) lo único que importa de $|P\rangle$ es su dirección. El multiplicar $|P\rangle$ por un número distinto de 0, no afecta a la cuestión de si satisface o no a (10).

Junto a la ecuación (10), consideraremos la ecuación imaginaria conjugada

$$\langle Q|\alpha = b\langle Q|, \tag{11}$$

en la que b es un número. Aquí las incógnitas son los números b y los bras $\langle Q|$ que no sean el cero. Dado que las ecuaciones (10) y (11) son de importancia primordial para la teoría, es conveniente dar un nombre especial al tipo de relación entre las cantidades a que dan lugar. Si una terna α, $|P\rangle$

y a verifican (10), diremos que a es un *autovalor (eigenvalue)* * del operador lineal α o de su correspondiente variable dinámica, y asimismo que $|P\rangle$ es un autoket *(eigenket)* del operador lineal o de la variable dinámica en cuestión. Además diremos que el autoket $|P\rangle$, *pertenece* al autovalor a. Análogamente, si una terna α, $\langle Q|$ y b verifican (11), diremos que b es un autovalor de α y que $\langle Q|$ es un autobra *(eigenbra)* perteneciente a dicho autovalor. Desde luego, las expresiones de autovalor, autoket y autobra *únicamente tienen sentido con respecto a un operador lineal o a una variable dinámica.*

Con esta terminología, si multiplicamos un autoket de α por un número distinto de cero, el ket resultante es un nuevo autoket perteneciente al mismo autovalor al que pertenecía el ket de partida. Pueden existir dos o más autokets de un operador lineal independientes que pertenezcan a un mismo autovalor; así por ejemplo, la ecuación (10) puede tener varias soluciones $|P1\rangle$, $|P2\rangle$, $|P3\rangle$, ... que pertenezcan a un mismo autovalor a, y que sean independientes. En tal caso es evidente que cualquier combinación lineal de dichos autokets también es un autoket que pertenece al mismo autovalor del operador lineal, es decir que

$$c_1|P1\rangle + c_2|P2\rangle + c_3|P3\rangle + \ldots$$

es también solución de (10), siendo c_1, c_2, c_3, ... números cualesquiera.

En el caso particular de que el operador lineal α de las ecuaciones (10) y (11) sea un número k, evidentemente todo ket $|P\rangle$ y todo bra $\langle Q|$ satisfacen dichas ecuaciones con a y b iguales a k. Por tanto, todo número, considerado como operador lineal, tiene un único autovalor, y todos los kets y todos los bras son autovectores suyos pertenecientes a dicho autovalor.

La teoría de autovalores y autovectores de un operador lineal α que no sea real, no tiene gran aplicación en mecánica cuántica. Así pues, a estos efectos, nos limitaremos en adelante a considerar operadores lineales reales. Poniendo en lugar de α el operador lineal real ξ, resultan en lugar de las ecuaciones (10) y (11),

$$\xi|P\rangle = a|P\rangle, \tag{12}$$

$$\langle Q|\xi = b\langle Q|. \tag{13}$$

Podemos deducir inmediatamente tres importantes resultados.

(1) *Los autovalores son todos reales.* Para demostrar que todo a que verifique (13) es real, multipliquemos (12) por el bra $\langle P|$ a la izquierda, y resulta

$$\langle P|\xi|P\rangle = a\langle P|P\rangle.$$

---

* Algunas veces se utiliza la palabra 'propio' *(proper)* en lugar de 'auto' *(eigen)*, pero esto no es del todo satisfactorio ya que las palabras 'propio' e 'impropio' se utilizan a menudo con otros significados. Por ejemplo, en §§ 15 y 46 se utilizan las expresiones 'función impropia' y 'energía propia'.

Ahora, con ayuda de (4), en la que substituimos $\langle B|$ *por* $\langle P|$ y α por el operador lineal real ξ, resulta que el número $\langle P|\xi|P\rangle$ tiene que ser real, y como en virtud de § 6 $\langle P|P\rangle$ es real y distinto de cero, deducimos que a es real. Análogamente, multiplicando (13) por $|Q\rangle$ a la derecha, podríamos demostrar que b también es real.

Dada una solución de (12) formemos la ecuación imaginaria conjugada, que es

$$\langle P|\xi = a\langle P|$$

ya que tanto ξ como a son reales. Esta ecuación, conjugada imaginaria de (12), nos proporciona una solución de (13) tomando $\langle Q| = \langle P|$ y $b = a$. De aquí se deduce que

(II) *Los autovalores asociados a los autokets son los mismos que los autovalores asociados a los autobras.* $\alpha|A\rangle = a|A\rangle$ $\langle A|\alpha = \langle A|a$

(III) *El imaginario conjugado de un autoket es un autobra que pertenece al mismo autovalor, y recíprocamente.* El último resultado nos permite decir con razón que el estado que corresponde a cualquier autoket o a su bra imaginario conjugado es un *autoestado* de la variable dinámica real ξ.

Los autovalores y los autovectores de algunas variables dinámicas reales se utilizan ampliamente en mecánica cuántica y, por tanto, es conveniente utilizar una notación sistemática para indicarlos. La que vamos a exponer a continuación es adecuada para la mayoría de las aplicaciones. Si ξ es una variable dinámica real, sus autovalores los designaremos por ξ', ξ'', ξ^r, etc. Así una letra sirve para designar una *variable dinámica real* o un *operador lineal real*, y la misma letra con primas o índices indica un *número*, que es un autovalor del operador que simboliza la letra. De este modo, un autovector puede especificarse con el autovalor a que pertenece. Así $|\xi'\rangle$ representa un autoket que pertenece al autovalor ξ' de la variable dinámica ξ. Cuando consideremos explícitamente un caso en que a un autovalor de una variable real le corresponda más de un autoket, los distinguiremos unos de otros mediante un nuevo símbolo, o incluso por más de uno. Por tanto, si tratamos con dos autokets que pertenezcan al mismo autovalor ξ', les designaremos $|\xi'1\rangle$ y $|\xi'2\rangle$.

TEOREMA. *Dos autovectores de una misma variable dinámica real que pertenezcan a distintos autovalores son ortogonales.*

Para demostrar el teorema, consideremos dos autokets $|\xi'\rangle$ y $|\xi''\rangle$ de una misma variable real ξ, que pertenezcan respectivamente a los autovalores ξ' y ξ''. Tendremos las ecuaciones

$$\xi|\xi'\rangle = \xi'|\xi'\rangle, \tag{14}$$

$$\xi|\xi''\rangle = \xi''|\xi''\rangle. \tag{15}$$

Tomando la ecuación imaginaria conjugada de (14) obtenemos

$$\langle\xi'|\xi = \xi'\langle\xi'|.$$

que multiplicada a la derecha por $|\xi''\rangle$ nos da

$$\langle\xi'|\xi|\xi''\rangle = \xi'\langle\xi'|\xi''\rangle$$

y multiplicando la (15) por $\langle\xi'|$ a la izquierda

$$\langle\xi'|\xi|\xi''\rangle = \xi''\langle\xi'|\xi''\rangle.$$

Restando una de otra obtenemos

$$(\xi' - \xi'')\langle\xi'|\xi''\rangle = 0, \tag{16}$$

que nos demuestra que si $\xi' \neq \xi''$, entonces $\langle\xi'|\xi''\rangle = 0$, o sea que los autovectores $|\xi'\rangle$ y $|\xi''\rangle$ son ortogonales. A este teorema se le da el nombre de *teorema de ortogonalidad.*

Hemos considerado propiedades de los autovalores y de los autovectores de un operador lineal real, pero no nos hemos preguntado si, dado un operador lineal real, existen autovalores y autóvectores, ni tampoco cómo se pueden hallar en caso de que existan. Esta cuestión, en general, es muy difícil de resolver. Sin embargo, existe un caso particular muy útil que se puede estudiar con facilidad, y es cuando el operador lineal real ξ en cuestión satisface una ecuación algebraica

$$\phi(\xi) \equiv \xi^n + a_1\xi^{n-1} + a_2\xi^{n-2} + \ldots + a_n = 0, \tag{17}$$

en la que los coeficientes a_i sean números. Esta ecuación significa, como es sabido, que al aplicar el operador lineal $\phi(\xi)$ a cualquier ket o a cualquier bra da como resultado cero.

Sea (17) la ecuación algebraica más sencilla que satisface ξ. Vamos a demostrar que

(α) El número de autovalores de ξ es n.

(β) Existen tantos autokets de ξ que todo ket puede expresarse linealmente en función de ellos.

La forma algebraica $\phi(\xi)$ puede ser descompuesta en n factores lineales así

$$\phi(\xi) \equiv (\xi - c_1)(\xi - c_2)(\xi - c_3)\ldots(\xi - c_n) \tag{18}$$

siendo los c_i ciertos números que no tienen por qué ser todos distintos. Esta descomposición puede llevarse a cabo tanto si ξ es una variable algebraica ordinaria como si es un operador, ya que en (18) no hay ningún elemento que no conmute con ξ. Sea $\chi_r(\xi)$ el cociente de dividir $\phi(\xi)$ por $(\xi - c_r)$, es decir

$$\phi(\xi) \equiv (\xi - c_r)\chi_r(\xi) \qquad (r = 1, 2, 3, \ldots, n).$$

Entonces, para todo ket $|P\rangle$ se tiene

$$(\xi - c_r)\chi_r(\xi)|P\rangle = \phi(\xi)|P\rangle = 0. \tag{19}$$

Pero $\chi_r(\xi)|P\rangle$ no puede ser nulo para todo ket $|P\rangle$, ya que entonces también

sería nulo $\chi_r(\xi)$, en cuyo caso ξ satisfaría una ecuación algebraica de grado $n-1$, lo que está en contra de la hipótesis de que (17) es la ecuación más sencilla que satisface ξ. Si elegimos $|P\rangle$ de forma que $\chi_r(\xi)|P\rangle$ no sea nulo, la ecuación (19) nos muestra que $\chi_r(\xi)|P\rangle$ es un autoket de ξ perteneciente al autovalor c_r. El argumento es válido para todo valor de r desde 1 hasta n y, por tanto, cada uno de los c_i es un autovalor de ξ. Además, ningún otro número puede ser autovalor de ξ, ya que si ξ' es un autovalor cualquiera, que pertenece al autoket $|\xi'\rangle$, se tendrá

$$\xi|\xi'\rangle = \xi'|\xi'\rangle$$

de donde podemos deducir

$$\phi(\xi)|\xi'\rangle = \phi(\xi')|\xi'\rangle,$$

y puesto que el primer miembro se anula, debe ser $\phi(\xi') = 0$.

Para completar la demostración de (α) hemos de ver que todos los c_i que figuran en (18) son distintos. Supongamos que no lo fueran, y que c_s estuviera repetido m veces, siendo $m > 1$. En tal caso $\phi(\xi)$ sería de la forma

$$\phi(\xi) \equiv (\xi - c_s)^m \theta(\xi),$$

donde $\theta(\xi)$ sería una función racional y entera de ξ. De (17) deducimos

$$(\xi - c_s)^m \theta(\xi)|A\rangle = 0 \tag{20}$$

que se verifica para todo ket $|A\rangle$. Puesto que c_s es un autovalor de ξ, tiene que ser real y, por tanto, $\xi - c_s$ es un operador lineal real. Pero la ecuación (20) es de la misma forma que la (8) con $\xi - c_s$ en lugar de ξ y $\theta(\xi)|A\rangle$ en lugar de $|P\rangle$. En virtud del teorema relacionado con dicha ecuación (8), podemos deducir

$$(\xi - c_s)\theta(\xi)|A\rangle = 0.$$

Y puesto que el ket $|A\rangle$ es arbitrario, resulta

$$(\xi - c_s)\theta(\xi) = 0,$$

que está en contradicción con la hipótesis de que (17) es la ecuación más sencilla que verifica ξ. Así pues, los c_i son todos distintos y queda demostrado (α).

Sea $\chi_r(c_r)$ el número que se obtiene al substituir ξ por c_r en la expresión algebraica de $\chi_r(\xi)$. Puesto que todos los c_i son distintos, $\chi_r(c_r)$ no puede ser nulo. Consideremos la expresión

$$\sum_r \frac{\chi_r(\xi)}{\chi_r(c_r)} - 1. \tag{21}$$

Si en ella substituimos ξ por c_s, todos los términos de la suma se anulan con excepción del término para el que $r = s$, pues para $r \neq s$, $\chi_r(\xi)$, contiene

el factor $(\xi - c_s)$, y el término $r = s$ vale 1 y se cancela el —1. Por tanto, la expresión (21) se anula cuando ponemos en lugar de ξ cualquiera de los n números $c_1, c_2, \ldots, c_n$, y al ser de grado $n - 1$ en ξ, debe ser idénticamente nula. Así pues, si aplicamos el operador lineal (21) a un ket arbitrario $|P\rangle$ e igualamos el resultado a cero, obtenemos

$$|P\rangle = \sum_r \frac{1}{\chi_r(c_r)} \chi_r(\xi)|P\rangle. \tag{22}$$

Según (19), cada término de esta suma que no sea nulo es un autoket de ξ. La ecuación (22) expresa un ket arbitrario $|P\rangle$ en función de autokets de ξ, quedando así demostrado (β).

Como ejemplo sencillo podemos considerar un operador lineal real σ que satisfaga a la ecuación

$$\sigma^2 = 1. \tag{23}$$

En este caso σ tiene los dos autovalores 1 y —1. Todo ket $|P\rangle$ puede expresarse así

$$|P\rangle = \tfrac{1}{2}(1 + \sigma)|P\rangle + \tfrac{1}{2}(1 - \sigma)|P\rangle.$$

Es fácil ver que los dos términos del segundo miembro, si no se anulan, son autokets de σ, que pertenecen respectivamente a los autovalores 1 y —1.

10. *Observables*

Hemos hecho un conjunto de hipótesis acerca de cómo se representan los estados y las variables dinámicas en la teoría matemática. Dichas hipótesis por sí solas no constituyen leyes de la naturaleza, pero en cuanto hagamos nuevas hipótesis acerca del significado físico de la teoría, aparecerán como tales. Las nuevas hipótesis tienen que establecer relaciones entre los resultados de observación por un lado, y las ecuaciones del formalismo matemático por el otro.

Toda observación consiste en medir una variable dinámica. Desde el punto de vista físico es evidente que el resultado de dicha medición debe ser siempre un número real, y así supondremos que cualquier variable dinámica que podamos medir debe ser una variable dinámica real. Puede pensarse que sería posible medir una variable dinámica compleja midiendo su parte real y su parte imaginaria por separado, pero esto llevaría consigo dos medidas o dos observaciones. Si bien en mecánica clásica es posible, en mecánica cuántica no lo es, pues en general las medidas interfieren entre sí — no se puede admitir que dos observaciones se realicen exactamente a la vez, y si se hacen una detrás de otra en rápida sucesión, generalmente la primera altera el estado del sistema e introduce una indeterminación que afecta a la segunda. Por lo tanto, debemos exigir que las

variables dinámicas que se pueden medir sean reales, y la condición para ello en mecánica cuántica es la de § 8. Sin embargo, no toda variable dinámica real puede ser medida. Como veremos más adelante, es necesaria otra restricción.

Hagamos ahora algunas hipótesis acerca de la interpretación física de la teoría. *Si el sistema dinámico está en un autoestado de una variable dinámica real* ξ *perteneciente al autovalor* ξ', *entonces estamos totalmente seguros de que el resultado de medir* ξ *es el número* ξ'. Recíprocamente, *si el sistema está en un estado tal que al medir una variable dinámica real* ξ *estamos completamente seguros de obtener un resultado particular determinado* (y no distintos posibles resultados según una ley probabilística, como ocurre en general), *entonces el estado es un autoestado de* ξ, *y el resultado de la medida es el autovalor de* ξ *a que pertenece dicho autoestado.* Dado que los autovalores de los operadores lineales reales son siempre números reales, estas hipótesis son viables.

Señalemos algunas consecuencias inmediatas de estas hipótesis. Si tenemos dos o más autoestados de una variable dinámica real ξ pertenecientes a un mismo autovalor ξ', todo estado formado por superposición de ellos también será un autoestado de ξ perteneciente al autovalor ξ'. De aquí se deduce que si el sistema está en un estado formado por superposición de otros varios para los que estamos completamente seguros de que al medir ξ en cada uno de ellos obtenemos un mismo resultado ξ', entonces al medir ξ en él obtenemos también con toda seguridad el resultado ξ'. Esto nos da cierta luz acerca del significado físico del principio de superposición de los estados. Otra consecuencia es que dos autoestados de ξ pertenecientes a distintos autovalores son ortogonales. De aquí se deduce que si tenemos dos estados de un sistema y estamos completamente seguros de que al medir ξ en cada uno de ellos obtenemos un único resultado, distinto para uno y otro, entonces los dos estados son ortogonales. Esto nos da, a su vez, alguna luz acerca del significado físico de la ortogonalidad de los estados.

Cuando medimos una variable dinámica real ξ, la alteración que lleva consigo el acto de la medida produce un cambio del estado del sistema dinámico. Si llevamos a cabo una segunda medición de la misma variable dinámica ξ inmediatamente después de la primera, por continuidad física el resultado debe ser el mismo de antes. Por tanto, después de haber realizado la primera medición no hay ninguna indeterminación en el resultado de la segunda, o sea que después de realizar la primera medición, el sistema está en un autoestado de la variable dinámica ξ, siendo el autovalor al que pertenece dicho estado el resultado de la primera medida. Esta conclusión tiene que continuar siendo válida aunque no llevemos a cabo la segunda medida. Así vemos que toda medida obliga al sistema a saltar a un autoestado de la variable dinámica medida, que además pertenece a un autovalor igual al resultado de la medida.

Podemos concluir que, cualquiera que sea el estado en que se encuentre

el sistema, *todo resultado de medir una variable dinámica real es uno de sus autovalores.* Y recíprocamente, *todo autovalor es un posible resultado de medida de la variable dinámica para algún estado del sistema,* ya que con toda probabilidad será el resultado de la medida cuando el estado del sistema sea un autoestado perteneciente a dicho autovalor. Todo esto nos muestra el significado físico de los autovalores. El conjunto de autovalores de una variable dinámica real constituye precisamente el espectro de los posibles resultados de medida de dicha variable dinámica, y por esta razón el cálculo de los autovalores constituye un problema importante.

Asimismo hacemos la hipótesis siguiente sobre la interpretación física de la teoría: *cuando medimos una cierta variable dinámica real ξ con el sistema en un estado determinado, los estados a los que puede saltar el sistema a causa de la medida son tales que el estado original es dependiente de ellos.* Pero como por otro lado los estados a los que puede saltar el sistema son todos autoestados de ξ, resulta que el estado original es dependiente de autoestados de ξ. Teniendo en cuenta que el estado original puede ser *cualquiera,* podemos concluir que todo estado es dependiente de autoestados de ξ. Si definimos un *conjunto completo* de estados como aquel para el que todo estado es dependiente de los estados del conjunto, nuestra conclusión también se puede formular así: los autoestados de ξ constituyen un conjunto completo.

No toda variable dinámica real tiene suficientes autoestados para que constituyan un conjunto completo. Pero aquellas cuyos autoestados no forman un conjunto completo no representan cantidades que se puedan medir. Así obtenemos una nueva condición que debe verificar toda variable dinámica, además de la de ser real, para que se pueda medir. Toda variable dinámica real cuyos autoestados constituyan un conjunto completo, la denominaremos con el nombre genérico de *observable.* Por tanto, toda cantidad que se pueda medir es un observable.

La cuestión que se plantea ahora es si todo observable puede ser medido. La respuesta teórica es que sí. En la práctica puede resultar muy difícil o incluso puede estar fuera del alcance del ingenio del experimentador, el imaginar un aparato capaz de medir un observable particular, pero según la teoría siempre existe.

Examinemos matemáticamente la condición para que una variable dinámica real ξ sea un observable. Sus autovalores pueden constituir un conjunto discreto (finito o infinito) de números o bien un conjunto formado por todos los números de un cierto dominio de medida no nula, como todos los números comprendidos entre a y b. En el primer caso, la condición de que todo estado sea dependiente de los autoestados de ξ es que todo ket pueda expresarse como combinación lineal de autokets de ξ. En el segundo caso la condición debe modificarse, ya que en lugar de la suma puede aparecer una integral; en este caso, todo ket $|P\rangle$ ha de poderse expresar como una integral de autokets de ξ.

$$|P\rangle = \int |\xi'\rangle d\xi', \tag{24}$$

donde $|\xi'\rangle$ son autokets de ξ pertenecientes al autovalor ξ', y la integral se extiende al dominio de los autovalores a los que pertenecen los autokets de ξ de los que depende dicho ket. Pero no todo ket que dependa de autokets de ξ puede expresarse en la forma que indica el segundo miembro de (24), pues por ejemplo uno de los mismos autokets no se puede expresar así, y más en general, ocurre lo mismo con cualquier suma de ellos. Por tanto, la condición de que los autoestados de ξ constituyan un conjunto completo tiene que formularse diciendo que todo ket $|P\rangle$ ha de poderse expresar como una integral más una suma de autokets de ξ, o sea

$$|P\rangle = \int |\xi' c\rangle d\xi' + \sum_r |\xi^r d\rangle, \tag{25}$$

Los $|\xi' c\rangle$ que figuran en la integral y los $|\xi^r d\rangle$ de la suma son autokets de ξ, y les hemos añadido las letras c y d respectivamente para indicar que no tienen por qué ser iguales aunque los autovalores ξ' y ξ^r coincidan. La integral se extiende a todo el dominio de valores, y la suma a un subconjunto discreto cualquiera. Si esta condición se satisface en el caso de que los autovalores de ξ constituyan un dominio de medida no nula, entonces ξ es un observable.

Puede darse un caso más general aún, cuando los autovalores de ξ son, además de todos los números de un cierto dominio de medida no nula, un conjunto discreto de números situado fuera de dicho dominio. En este caso, la condición de que ξ sea un observable sigue siendo que todo ket pueda expresarse en la forma que indica el segundo miembro de (25), pero advirtiendo que ahora la suma respecto de r se extiende además de a un subconjunto discreto cualquiera de los valores del dominio como, antes, a los números del conjunto discreto situado fuera de dicho dominio.

A menudo resulta muy difícil establecer desde el punto de vista matemático si una variable dinámica real concreta satisface o no la condición para ser un observable, ya que, en general, el problema de hallar los autovalores y los autovectores es muy difícil de resolver. Sin embargo, pueden existir poderosas razones experimentales que nos permitan admitir que dicha variable dinámica puede ser medida, y en tal caso supondremos lógicamente que es un observable pese a carecer de una demostración matemática. Esto lo haremos con frecuencia durante el desarrollo de la teoría, y así supondremos por ejemplo, que la energía de cualquier sistema dinámico es siempre un observable, si bien con los métodos actuales del análisis matemático sólo se ha podido demostrar en casos muy sencillos.

En el caso particular de que la variable dinámica sea un número, todo estado es autoestado de él, y evidentemente dicha variable dinámica es un observable. Al medirlo obtendremos siempre el mismo resultado, o sea que se trata de una constante física, como por ejemplo la carga del electrón.

Por tanto, una constante física se puede considerar en mecánica cuántica bien como un observable con un único autovalor o como un simple número que aparece en las ecuaciones, siendo ambos puntos de vista completamente equivalentes.

Si la variable dinámica real satisface una ecuación algebraica, entonces el resultado (β) de la sección anterior nos dice que la variable dinámica en cuestión es un observable. Un observable de este tipo tiene un número finito de autovalores. Recíprocamente, todo observable que tenga un número finito de autovalores satisface una ecuación algebraica, pues si el observable ξ tiene como autovalores ξ', ξ'', ..., ξ^n, entonces se verifica

$$(\xi-\xi')(\xi-\xi'')\ldots(\xi-\xi^n)|P\rangle = 0$$

para todo $|P\rangle$ que sea autoket de ξ y, por lo tanto, también para cualquier $|P\rangle$, pues como ξ es un observable todo ket puede expresarse como suma de autokets de ξ, luego

$$(\xi-\xi')(\xi-\xi'')\ldots(\xi-\xi^n) = 0. \tag{26}$$

Como ejemplo tomemos el operador lineal $|A\rangle\langle A|$, siendo $|A\rangle$ un ket normalizado. Según (7), dicho operador es real y su cuadrado vale

$$\{|A\rangle\langle A|\}^2 = |A\rangle\langle A|A\rangle\langle A| = |A\rangle\langle A| \tag{27}$$

ya que $\langle A|A\rangle = 1$. Es decir, que su cuadrado es igual a sí mismo y, por lo tanto, satisface una ecuación algebraica y, en consecuencia, es un observable. Sus autovalores son 1 y 0; $|A\rangle$ es el autoket que pertenece al autovalor 1 y todos los kets ortogonales a $|A\rangle$ son autokets que pertenecen al autovalor 0. La medición de dicho observable dará con toda certeza 1 cuando el sistema dinámico esté en el estado correspondiente a $|A\rangle$, y 0 cuando esté en un estado cualquiera ortogonal a éste. Así pues, el observable representa la cantidad que determina si el sistema está en el estado $|A\rangle$ o no.

Antes de concluir esta sección vamos a examinar las condiciones que se necesitan para que una integral como la de (24) tenga sentido. Sean $|X\rangle$ e $|Y\rangle$ dos kets que puedan expresarse como integrales de autokets del observable ξ,

$$|X\rangle = \int |\xi' x\rangle d\xi', \qquad |Y\rangle = \int |\xi'' y\rangle d\xi'',$$

donde x e y sirven para distinguir ambos integrandos. En este caso, tomando la ecuación imaginaria conjugada de la primera y multiplicándola por la segunda tenemos

$$\langle X|Y\rangle = \int\int \langle \xi' x|\xi'' y\rangle d\xi' d\xi''. \tag{28}$$

Consideremos la integral

$$\int \langle \xi' x|\xi'' y\rangle d\xi''. \tag{29}$$

Según el teorema de ortogonalidad, el integrando se anula en todo el inter-

valo de integración excepto en el punto $\xi'' = \xi'$. Si el integrando fuera finito en dicho punto, la integral (29) se anularía, y si ello ocurriese para todo ξ', según (28) resultaría $\langle X|Y\rangle$ igual a cero. Pero como en general $\langle X|Y\rangle$ no es nulo, $\langle \xi' x|\xi' y\rangle$ ha de ser en general infinito, de tal manera que (29) resulte distinto de cero y finito. El tipo de infinito que esto exige se discutirá en § 15.

Implícitamente habíamos supuesto hasta ahora que nuestros bras y kets tenían longitud finita y que sus productos escalares eran finitos. Ahora nos vemos obligados a limitar dicha condición cuando tratamos con autovectores de un observable cuyos autovalores forman un dominio de medida no nula. Si no la limitásemos no sería posible considerar casos en que los autovalores constituyeran un dominio de medida no nula, y nuestra teoría resultaría insuficiente para una gran parte de los problemas reales.

Tomando $|Y\rangle = |X\rangle$ en (28), resulta que en general $\langle \xi' x|\xi' x\rangle$ es infinito. Para $|\xi' x\rangle \neq 0$ tomaremos

$$\int \langle \xi' x|\xi'' x\rangle \, d\xi'' > 0, \tag{30}$$

como el axioma correspondiente a (8) de § 6 para los vectores de longitud infinita.

El espacio de los bras o el de los kets, cuando exigimos que los vectores tengan longitud finita y que los productos escalares sean finitos, constituye lo que los matemáticos denominan un *espacio de Hilbert*. Los bras y los kets que empleamos aquí dan lugar a un espacio más general de lo que corresponde a un espacio de Hilbert.

Veamos ahora que, si exigimos que no haya más que un término en la suma relativo a un mismo autovalor, la descomposición de un ket $|P\rangle$ en la forma señalada en (25) es única. Para demostrarlo supongamos que, por el contrario, existieran dos posibles descomposiciones distintas de $|P\rangle$. Restando una de la otra, obtendríamos una ecuación de la forma

$$0 = \int |\xi' a\rangle \, d\xi' + \sum_s |\xi^s b\rangle, \tag{31}$$

en la que a y b son nuevos símbolos para distinguir los autovectores resultantes, y donde el subíndice s incluye todos los términos que hayan quedado después de restar. Si en la suma de (31) hubiese un término relativo a un autovalor ξ^t que no forme parte del dominio continuo, multiplicando dicha ecuación por $\langle \xi^t b|$ a la izquierda y aplicando el teorema de ortogonalidad, resultaría

$$0 = \langle \xi^t b|\xi^t b\rangle,$$

lo que está en contra de la hipótesis (8) de § 6. Asimismo, si el integrando de (31) fuera distinto de cero para algún autovalor ξ'' que no figurase en la

suma para ningún valor de s, multiplicando (31) por $\langle \xi'' a|$ a la izquierda y aplicando el teorema de ortogonalidad, resultaría

$$0 = \int \langle \xi'' a | \xi' a \rangle \, d\xi',$$

lo que va en contra de (30). Por último, si en la suma de (31) existiera un término relativo a un valor propio ξ^t que figurase en el dominio de integración, multiplicando la ecuación por $\langle \xi^t b|$ a la izquierda, resultaría

$$0 = \int \langle \xi^t b | \xi' a \rangle \, d\xi' + \langle \xi^t b | \xi^t b \rangle \tag{32}$$

y multiplicando por $\langle \xi^t a|$ a la izquierda

$$0 = \int \langle \xi^t a | \xi' a \rangle \, d\xi' + \langle \xi^t a | \xi^t b \rangle. \tag{33}$$

Pero la integral de (33) es finita, luego, también deben serlo $\langle \xi^t a|\xi^t b\rangle$ y $\langle \xi^t b|\xi^t a\rangle$. Esto implica que la integral que figura en (32) sea nula y, por lo tanto, $\langle \xi^t b|\xi^t b\rangle$ también, resultando una nueva contradicción. Así pues, todos los términos de (31) son nulos, y la descomposición de un ket $|P\rangle$ en la forma indicada por el segundo miembro de (25) es única.

11. *Funciones de observables*

Sea ξ un observable. Podemos multiplicarlo por cualquier número real k y obtener otro nuevo observable $k\xi$. Para que nuestra teoría sea coherente es necesario que cuando el sistema se halle en un estado en el que con toda certeza obtenemos el resultado ξ' al medir el observable ξ, el resultado de medir el observable $k\xi$ sea con toda certeza $k\xi'$. Es fácil comprobar que se satisface dicha condición. El ket que corresponde a un estado en el que con toda certeza al medir ξ obtenemos el resultado ξ' es un autoket de ξ al que llamaremos $|\xi'\rangle$, que verifica

$$\xi|\xi'\rangle = \xi'|\xi'\rangle.$$

De esta ecuación resulta también

$$k\xi|\xi'\rangle = k\xi'|\xi'\rangle,$$

que nos indica que $|\xi'\rangle$ es un autoket de $k\xi$ perteneciente al autovalor $k\xi'$, y que, por lo tanto, al medir $k\xi$ obtendremos con toda seguridad el resultado $k\xi'$.

Más en general, podemos tomar cualquier función real de ξ, por ejemplo $f(\xi)$, y considerarla como un nuevo observable que queda medido automáticamente cuando se mide ξ, pues una medida experimental de ξ también

nos permite disponer del valor de $f(\xi)$. No es necesario que $f(\xi)$ sea real, y cuando no lo es tanto su parte real como su parte imaginaria son observables que quedan determinados automáticamente cuando medimos ξ. Para que la teoría sea coherente, es necesario que al medir la parte real y la parte imaginaria de $f(\xi)$ en un estado en el que una medida de ξ dé con certeza el resultado ξ', obtengamos con toda seguridad las partes real e imaginaria de $f(\xi')$. Si $f(\xi)$ es desarrollable en serie de potencias

$$f(\xi) = c_0 + c_1\xi + c_2\xi^2 + c_3\xi^3 + \ldots.$$

siendo los c_i ciertos números, la condición dada se puede comprobar con los métodos del álgebra elemental. Para funciones f más generales no será posible comprobarla. En este caso puede utilizarse dicha condición para definir $f(\xi)$, que aún no habíamos definido desde el punto de vista matemático. De este modo podemos dar una definición de función de un observable más general que la que se obtiene mediante series de potencias.

Definiremos $f(\xi)$ como el operador lineal que verifica

$$f(\xi)|\xi'\rangle = f(\xi')|\xi'\rangle \tag{34}$$

para todo autoket $|\xi'\rangle$ de ξ, siendo $f(\xi')$ una función numérica de ξ'. Puede verse fácilmente que la definición dada está justificada cuando se aplica a autokets $|\xi'\rangle$ que no sean todos independientes, pues si tenemos un autoket $|\xi'A\rangle$ que dependa de otros autokets de ξ, todos ellos tienen que pertenecer al mismo autovalor ξ', ya que de no ser así llegaríamos a una ecuación análoga a la (31) que según hemos visto es imposible. Y así, multiplicando la ecuación que expresa $|\xi'A\rangle$ como combinación lineal de los otros autokets de ξ por $f(\xi)$ a la izquierda, no hacemos más que multiplicar cada término por el mismo número $f(\xi')$, obteniéndose con ello una ecuación coherente. Además, la ecuación (34) es suficiente para definir completamente el operador lineal $f(\xi)$, pues para obtener el resultado de aplicar $f(\xi)$ a un ket $|P\rangle$ arbitrario basta desarrollar $|P\rangle$ en la forma indicada por el segundo miembro de (25), y tomar

$$f(\xi)|P\rangle = \int f(\xi')|\xi'c\rangle\, d\xi' + \sum_r f(\xi^r)|\xi^r d\rangle. \tag{35}$$

El complejo conjugado $\overline{f(\xi)}$ de $f(\xi)$ queda definido por la ecuación imaginaria conjugada de la (34), o sea

$$\langle\xi'|\overline{f(\xi)} = \bar{f}(\xi')\langle\xi'|,$$

para todo autobra $\langle\xi'|$, siendo $\bar{f}(\xi')$ la función compleja conjugada de $f(\xi')$. Si sustituimos ξ' por ξ'' en esta ecuación, y la multiplicamos a la derecha por el ket arbitrario $|P\rangle$, con ayuda del desarrollo (25) y del teorema de ortogonalidad, resulta

$$\begin{aligned}\langle\xi''|\overline{f(\xi)}|P\rangle &= \bar{f}(\xi'')\langle\xi''|P\rangle \\ &= \int \bar{f}(\xi'')\langle\xi''|\xi' c\rangle\, d\xi' + \sum_r \bar{f}(\xi'')\langle\xi''|\xi' d\rangle \\ &= \int \bar{f}(\xi'')\langle\xi''|\xi' c\rangle\, d\xi' + \bar{f}(\xi'')\langle\xi''|\xi'' d\rangle\end{aligned} \tag{36}$$

en donde $\langle\xi''|\xi'' d\rangle$ es igual a cero si ξ'' no es ninguno de los autovalores que figuran en la suma (25). Poniendo ahora en (35) en lugar de $f(\xi')$ la función compleja conjugada $\bar{f}(\xi')$ y multiplicando a la izquierda por $\langle\xi''|$, resulta

$$\langle\xi''|\bar{f}(\xi)|P\rangle = \int \bar{f}(\xi')\langle\xi''|\xi' c\rangle\, d\xi' + \bar{f}(\xi'')\langle\xi''|\xi'' d\rangle.$$

El segundo miembro de esta ecuación es igual que el de (36), pues los integrandos son nulos para todo $\xi' \neq \xi''$, y por tanto,

$$\langle\xi''|\overline{f(\xi)}|P\rangle = \langle\xi''|\bar{f}(\xi)|P\rangle.$$

Esta ecuación es válida para todo autobra $\langle\xi''|$ y para todo ket $|P\rangle$ de modo que

$$\bar{f}(\xi) = \bar{f}(\xi), \tag{37}$$

Por tanto, *el complejo conjugado del operador lineal $f(\xi)$ es igual a la función $\bar{f}$ de ξ, compleja conjugada de f.*

Como corolario resulta que si $f(\xi')$ es una función real de ξ', $f(\xi)$ es un operador lineal real. Por tanto, $f(\xi)$ también es un observable, ya que además todo autoestado de ξ también es autoestado de $f(\xi)$ y, en consecuencia, sus autoestados constituyen un conjunto completo.

Con la definición dada, *estamos en condiciones de asignar un significado a cualquier función f de un observable, cuya función de variable real $f(x)$ correspondiente tenga un dominio de existencia que abarque todos los autovalores del observable.* Si el dominio de existencia de dicha función contiene otros puntos que no correspondan a autovalores, los valores de $f(x)$ en ellos no afectan para nada a la función del observable. La función no tiene por qué ser ni analítica ni continua. Los autovalores de la función f de un observable son iguales a la función f de los autovalores de dicho observable.

Es importante tener en cuenta que para poder definir la función f de un observable es necesario que para cada x que sea autovalor del observable exista un único valor de $f(x)$. Es decir, la función $f(x)$ debe ser uniforme. Podemos poner de manifiesto este hecho haciendo la siguiente consideración: cuando tenemos un observable $f(A)$ que es función real del observable A, ¿es a su vez A función del observable $f(A)$? La respuesta es que sí cuando a cada autovalor A' de A le corresponda un valor distinto de $f(A')$; pero si existen dos autovalores de A, por ejemplo A' y A'', para los que se tenga $f(A') = f(A'')$, entonces no existiría un único autovalor de A que

correspondiera al autovalor $f(A')$ del observable $f(A)$ y, por lo tanto, A no sería función del observable $f(A)$.

Es fácil comprobar desde el punto de vista matemático, y a partir de la definición, que la suma o el producto de dos funciones de un observable es otra función del mismo, y que una función de una función de un observable también es función de dicho observable. Asimismo es inmediato comprobar que toda la teoría de funciones de un observable es simétrica respecto a bras y kets, y que hubiéramos podido proceder del mismo modo con la ecuación

$$\boxed{\langle\xi'|f(\xi) = f(\xi')\langle\xi'|} \tag{38}$$

en lugar de la (34).

Terminaremos esta sección examinando dos ejemplos de gran interés práctico: la función recíproca y la raíz cuadrada. El recíproco de un observable existe siempre que el observable no tenga el autovalor cero. Si el observable α no tiene el autovalor cero, su recíproco, que lo representaremos por α^{-1} o $1/\alpha$ verificará

$$\alpha^{-1}|\alpha'\rangle = \alpha'^{-1}|\alpha'\rangle, \tag{39}$$

donde $|\alpha'\rangle$ es un autoket de α que pertenece al autovalor α'. Por tanto,

$$\alpha\alpha^{-1}|\alpha'\rangle = \alpha\alpha'^{-1}|\alpha'\rangle = |\alpha'\rangle.$$

Puesto que esta relación es válida para todo autoket $|\alpha'\rangle$, debe verificarse

$$\alpha\alpha^{-1} = 1. \tag{40}$$

Y análogamente,

$$\alpha^{-1}\alpha = 1. \tag{41}$$

Cualquiera de estas dos ecuaciones es suficiente para definir completamente α^{-1}, siempre que α no tenga el autovalor cero. En efecto, en el caso de la ecuación (40), sea x un operador lineal cualquiera que verifique la ecuación

$$\alpha x = 1.$$

Multipliquemos ambos miembros a la izquierda por el operador α^{-1} definido en (39). El resultado es

$$\alpha^{-1}\alpha x = \alpha^{-1}$$

y según (41)

$$x = \alpha^{-1}.$$

Las ecuaciones (40) y (41) se pueden utilizar para definir el recíproco (si es que existe) de un operador lineal α más general que no tiene por qué ser real. En este caso, puede que no sea suficiente una sola de las ecuaciones. Si dos operadores lineales α y β tienen recíprocos, su producto α y β tiene por recíproco

$$(\alpha\beta)^{-1} = \beta^{-1}\alpha^{-1}, \tag{42}$$

$(\alpha\beta)^{\dagger} = \beta^{\dagger}\alpha^{\dagger}$

que se obtiene tomando el recíproco de cada factor e invirtiendo el orden. Podemos comprobar (42) sin más que multiplicar el segundo miembro por $\alpha\beta$ a la derecha o a la iquierda y ver que se obtiene la identidad. Esta regla para obtener el recíproco de un producto puede generalizarse inmediatamente para más de dos factores, y así

$$(\alpha\beta\gamma\ldots)^{-1} = \ldots\gamma^{-1}\beta^{-1}\alpha^{-1}.$$

La raíz cuadrada de un observable α existe siempre, y cuando α no tiene autovalores negativos, es real. La representaremos por $\sqrt{\alpha}$ o por $\alpha^{\frac{1}{2}}$. Verifica la relación

$$\sqrt{\alpha}|\alpha'\rangle = \pm\sqrt{\alpha'}|\alpha'\rangle \tag{43}$$

para todo autoket $|\alpha'\rangle$ de α perteneciente a un autovalor α'. Por tanto,

$$\sqrt{\alpha}\sqrt{\alpha}|\alpha'\rangle = \sqrt{\alpha'}\sqrt{\alpha'}|\alpha'\rangle = \alpha'|\alpha'\rangle = \alpha|\alpha'\rangle,$$

y como esto es válido para todo autoket $|\alpha'\rangle$, resulta

$$\sqrt{\alpha}\sqrt{\alpha} = \alpha \tag{44}$$

Dada la ambigüedad de signo en (43), existirán muchas determinaciones posibles de la raíz cuadrada. Para fijar una de ellas tenemos que especificar el signo que corresponde a cada autovalor. Dicho signo puede cambiar irregularmente de un autovalor a otro, pues en cualquier caso (43) define un operador lineal $\sqrt{\alpha}$ que verifica (44) y que constituye, por tanto, una raíz cuadrada de α. Si existe algún autovalor de α al que le pertenezcan dos o más autokets independientes, entonces según la definición de función de un observable, hemos de atribuir el mismo signo a todos ellos en (43). Sin embargo, si tomásemos distintos signos, la ecuación (44) continuaría siendo válida y por tanto, salvo en el caso particular de que a cada autovalor de α le corresponda un único autoket de α independiente, la ecuación (44) por sí sola no es suficiente para definir $\sqrt{\alpha}$.

El número total de determinaciones posibles de la raíz cuadrada de un observable es 2^n, siendo n el número total de autovalores no nulos. En la práctica únicamente se utiliza la función raíz cuadrada para observables que no tienen autovalores negativos, y la determinación de la raíz que tiene utilidad es la que se obtiene tomando siempre el signo positivo en (43). Tal determinación se denomina *la raíz cuadrada positiva.*

12. *Interpretación física general*

Las hipótesis introducidas al principio de § 10 nos dan una interpretación física de la teoría matemática que sólo se puede aplicar en casos particulares, pues únicamente sirve para autoestados. Necesitamos alguna

(*) la cosa [illegible] complejo [illegible] (T. Regge 1958)

hipótesis más general que también nos permita obtener información física a partir de la teoría matemática cuando no consideremos autoestados.

En mecánica clásica decimos que un observable 'tiene un valor' para cualquier estado del sistema. ¿Qué es lo que corresponde a esto en mecánica cuántica? Si consideramos un observable ξ y dos estados x e y correspondientes a los dos vectores $\langle x|$ e $|y\rangle$, podemos formar el número $\langle x|\xi|y\rangle$. Dicho número no tiene características análogas al valor que 'tiene' un observable en la teoría clásica por tres razones principales: (I) hace referencia a *dos* estados, mientras que el valor clásico siempre se refiere a *uno*, (II) en general no es un número real, y (III) no está unívocamente determinado por el observable y los estados, pues los vectores $\langle x|$ e $|y\rangle$ contienen factores numéricos arbitrarios. Incluso si imponemos la condición de que $\langle x|$ e $|y\rangle$ estén normalizados, $\langle x|\xi|y\rangle$ continuará estando indeterminado en un factor de módulo uno. Sin embargo, si tomamos dos estados idénticos y elegimos $|y\rangle$ de modo que sea el vector imaginario conjugado de $\langle x|$, ninguna de las tres razones apuntadas tiene validez. El número que obtenemos en este caso $\langle x|\xi|x\rangle$ es necesariamente real y está unívocamente determinado si $\langle x|$ está normalizado, pues si multiplicamos $\langle x|$ por el factor numérico e^{ic}, siendo c un número real, tenemos que multiplicar $|x\rangle$ por e^{-ic}, y $\langle x|\xi|x\rangle$ toma el mismo valor de antes.

Por tanto, podríamos hacer la sugestiva hipótesis de que cuando el sistema está en el estado x, el observable ξ 'tiene un valor' $\langle x|\xi|x\rangle$ en un sentido análogo al clásico. Sin embargo, esta hipótesis no sería satisfactoria por la siguiente razón. Consideremos un segundo observable η que según la hipótesis anterior tendría un valor $\langle x|\eta|x\rangle$ cuando el sistema está en el mismo estado de antes. Por la analogía clásica, para este estado la suma de los dos observables debería tener un valor igual a la suma de los valores de los dos observables por separado, y el producto un valor igual al producto de dichos valores. La hipótesis nos llevaría al valor $\langle x|\xi + \eta|x\rangle$ para la suma de los dos observables, que coincide con la suma de $\langle x|\xi|x\rangle$ y $\langle x|\eta|x\rangle$, pero para el producto nos conduciría al valor $\langle x|\xi\eta|x\rangle$ o bien al $\langle x|\eta\xi|x\rangle$, y ninguno de ellos está relacionado de forma simple con $\langle x|\xi|x\rangle$ y $\langle x|\eta|x\rangle$.

Pero como la hipótesis sólo falla para el producto, sería lógico decir que $\langle x|\xi|x\rangle$ es el *valor medio* (que también se llama valor esperado) del observable ξ en el estado x, ya que el promedio de la suma de dos cantidades tiene que ser igual a la suma de promedios de dichas cantidades, pero el promedio del producto no tiene por qué ser igual al producto de promedios. Por tanto, hacemos la hipótesis general de que *si medimos el observable ξ un gran número de veces cuando el sistema está en el estado que corresponde a $|x\rangle$, el valor medio de todos los resultados obtenidos será $\langle x|\xi|x\rangle$ si $|x\rangle$ está normalizado.* Si $|x\rangle$ no está normalizado, como ocurrirá necesariamente en el caso de que el estado x sea un autoestado de cierto observable perteneciente a un autovalor que forma parte de un dominio de autovalores de medida no nula, la hipótesis se convierte en que el valor medio de las medi-

das de ξ es proporcional a $\langle x|\xi|x\rangle$. Esta hipótesis general nos da la base para la interpretación física general de la teoría.

Únicamente podemos decir en mecánica cuántica que un observable 'tiene un valor' particular en un estado determinado, cuando estemos totalmente seguros de que el resultado de medirlo es siempre dicho valor particular, o sea, cuando el sistema esté en un autoestado del observable. Si convenimos en utilizar la expresión de que un observable 'tiene un valor' con este significado restringido, es fácil comprobar con la ayuda del álgebra introducida, que si dos observables tienen un valor para un estado particular, entonces para dicho estado la suma de los dos observables (si es que es un observable *) tiene un valor igual a la suma de los valores de los dos observables, y el producto de ellos (si es que es un observable **) tiene un valor igual al producto de los valores de los dos observables.

En el caso general no tiene sentido hablar del valor que tiene un observable para un estado particular del sistema, pero sí podemos hablar del valor medio de dicho observable en ese estado. Podemos incluso ir más allá y hablar de la probabilidad de que al medirlo obtengamos dicho valor. Esta probabilidad puede obtenerse según la hipótesis general del siguiente modo.

Sea el observable ξ y consideremos el estado correspondiente al ket normalizado $|x\rangle$. La hipótesis general no sólo nos da el valor medio de ξ, que es $\langle x|\xi|x\rangle$, sino también el valor medio de cualquier función de ξ tal como $f(\xi)$, que será $\langle x|f(\xi)|x\rangle$. Tomemos como $f(\xi)$ una función de ξ que valga 1 cuando $\xi = a$, siendo a cualquier número real, y cero en cualquier otro caso. De acuerdo con nuestra teoría general de funciones de un observable, dicha función de ξ tiene sentido, y podemos designarla $\delta_{\xi a}$ en conformidad con la notación general del símbolo δ con dos subíndices dada en la pág. 73 (ecuación 17)). El valor medio de dicha función de ξ es precisamente la probabilidad P_a de que ξ tenga el valor a. O sea

$$P_a = \langle x|\delta_{\xi a}|x\rangle. \tag{45}$$

Si a no es autovalor de ξ, $\delta_{\xi a}$ multiplicado por cualquier autoket de ξ es cero y, por tanto, $\delta_{\xi a} = 0$ y $P_a = 0$. Esto está de acuerdo con la conclusión de § 10, de que todo resultado de la medida de un observable tiene que ser uno de sus autovalores.

Si los posibles resultados de medir ξ constituyen un conjunto de números de medida no nula, la probabilidad de que ξ tenga exactamente un determinado valor será nula para muchos problemas físicos. La magnitud que

---

* La suma de los dos observables no tiene por qué ser un observable, pues pudiera no tener suficientes autoestados para constituir un conjunto completo, en cuyo caso el operador suma considerado aisladamente no sería medible.

** Para el producto, la condición de que sea un observable puede no verificarse tanto porque sus autoestados no constituyan un conjunto completo, como por no ser real.

tiene importancia física en este caso es la probabilidad de que ξ tenga un valor perteneciente a un pequeño intervalo entre a y $a + da$. Esta probabilidad, que podemos simbolizar por $P(a)da$, es igual al valor esperado de la función de ξ que vale uno para valores de ξ comprendidos enrte a y $a + da$ y cero fuera de dicho intervalo. Según la teoría de funciones de un observable, esta función de ξ tiene sentido. Designándola por $\chi(\xi)$, tenemos

$$P(a)\, da = \langle x|\chi(\xi)|x\rangle. \tag{46}$$

Si el intervalo $(a, a + da)$ no contiene ningún autovalor de ξ, resulta como en el caso anterior, $\chi(\xi) = 0$ y $P(a) = 0$. Si $|x\rangle$ no está normalizado, los segundos miembros de (45) y (46) serán respectivamente proporcionales a las probabilidades de que ξ tenga un valor a y de que ξ tenga un valor comprendido entre a y $a + da$.

La hipótesis de § 10 de que cuando el sistema está en un autoestado de ξ perteneciente al autovalor ξ' toda medida de ξ da con seguridad el resultado ξ', está de acuerdo con la hipótesis general relativa a la interpretación física, y de hecho es consecuencia de ella. Según la hipótesis general, si $|\xi'\rangle$ es un autoket de ξ perteneciente al autovalor ξ', entonces en el caso de que los autovalores de ξ constituyan un conjunto discreto, resulta

$$\delta_{\xi a}|\xi'\rangle = 0 \quad \text{salvo para} \quad a = \xi',$$

y en el caso de que los autovalores de ξ constituyan un dominio de medida no nula

$$\chi(\xi)|\xi'\rangle = 0 \quad \text{salvo si el intervalo } (a, a + da) \text{ incluye a } \xi'.$$

En ambos casos, la probabilidad de que ξ tenga un valor distinto de ξ' cuando el sistema está en el estado correspondiente a $|\xi'\rangle$ es cero.

En la práctica es imposible obtener un sistema estrictamente en un autoestado de ξ perteneciente a un autovalor ξ' que forme parte de un intervalo de autovalores, ya que para ello sería necesaria una precisión absoluta. Lo máximo que se puede conseguir en la práctica es que ξ tenga un valor perteneciente a un pequeño intervalo alrededor del valor ξ'. El sistema estará entonces en un estado parecido a un autoestado de ξ. Así pues un autoestado perteneciente a un autovalor que forme parte de un intervalo de autovalores es una idealización matemática, y no se puede realizar en la práctica. A pesar de ello dichos autoestados tienen un papel muy importante en la teoría, y sin ellos no podríamos hacer gran cosa. La ciencia contiene muchos ejemplos de conceptos teóricos que son límites de cosas que se dan en la práctica, y que son de gran utilidad para la formulación precisa de las leyes de la naturaleza, pese a que no se pueden llevar a cabo experimentalmente; y éste no es más que uno de ellos. Es posible que el que la longitud de los kets correspondientes a estos estados sea infinita esté relacionado con su irrealizabilidad práctica, y que todos los estados realizables corresponden a kets que puedan ser normalizados y que, en consecuencia, forman parte de un espacio de Hilbert.

13. *Conmutabilidad y compatibilidad*

Un estado puede ser autoestado de dos observables a la vez. Supongamos que para los observables ξ y η exista un estado como éste correspondiente al ket $|A\rangle$. Entonces podremos escribir las ecuaciones

$$\xi|A\rangle = \xi'|A\rangle,$$

$$\eta|A\rangle = \eta'|A\rangle,$$

donde ξ' y η' son los autovalores respectivos de ξ y η. De aquí se deduce

$$\xi\eta|A\rangle = \xi\eta'|A\rangle = \xi'\eta'|A\rangle = \xi'\eta|A\rangle = \eta\xi'|A\rangle = \eta\xi|A\rangle,$$

o sea

$$(\xi\eta - \eta\xi)|A\rangle = 0.$$

Este resultado sugiere que la posibilidad de que existan autoestados comunes se verá favorecida cuando $\xi\eta - \eta\xi = 0$, es decir, cuando los dos observables conmuten. Si no conmutan, no es que no exista dicha posibilidad, pero es más bien excepcional. En cambio, *si conmutan* vamos a ver que *existen suficientes kets, que son a la vez autokets de uno y otro, para formar un conjunto completo.*

Sean ξ y η dos observables que conmutan. Tomemos un autoket $|\eta'\rangle$ de η, que pertenezca al autovalor η' y expresémoslo en función de autokets de ξ en la forma que indica el segundo miembro de (25)

$$|\eta'\rangle = \int |\xi'\eta' c\rangle\, d\xi' + \sum_r |\xi^r\eta' d\rangle. \tag{47}$$

Los autokets de ξ que figuran en el segundo miembro tienen el símbolo extra η' para indicar que provienen del desarrollo del ket $|\eta'\rangle$ y no de un ket cualquiera como ocurría en la ecuación (25). Veamos que cada uno de estos autokets de ξ es también autoket de η perteneciente al autovalor η'. Se tiene

$$0 = (\eta - \eta')|\eta'\rangle = \int (\eta - \eta')\, |\xi'\eta' c\rangle\, d\xi' + \sum_r (\eta - \eta')|\xi^r\eta' d\rangle. \tag{48}$$

Pero el ket $(\eta - \eta')|\xi^r\eta' d\rangle$ verifica

$$\begin{aligned}\xi(\eta - \eta')|\xi^r\eta' d\rangle = (\eta - \eta')\,\xi|\xi^r\eta' d\rangle &= (\eta - \eta')\xi^r|\xi^r\eta' d\rangle \\ &= \xi^r(\eta - \eta')|\xi^r\eta' d\rangle,\end{aligned}$$

lo que demuestra que es un autoket de ξ que pertenece al autovalor ξ^r; análogamente, el ket $(\eta - \eta')|\xi'\eta' c\rangle$ también es autoket de ξ y pertenece al autovalor ξ'. La ecuación (48) está constituida por una integral más una suma de autokets de ξ igualada a cero. En virtud del argumento utili-

zado para la ecuación (31), ello es imposible a menos que el integrando y cada uno de los términos de la suma sean nulos. Luego

$$(\eta-\eta')|\xi'\eta'c\rangle = 0, \qquad (\eta-\eta')|\xi^r\eta'd\rangle = 0,$$

lo que nos dice que todos los kets que aparecen en el segundo miembro de (47) son autokets tanto de ξ como de η. La ecuación (47) expresa $|\eta'\rangle$ en función de autokets comunes a ξ y η; y puesto que todo ket puede expresarse en función de autokets $|\eta'\rangle$ de η, resulta que todo ket puede expresarse en función de autokets comunes a ξ y a η y, en consecuencia, dichos kets constituyen un conjunto completo.

Los anteriores autokets $|\xi'\eta'c\rangle$ y $|\xi^r\eta'd\rangle$ comunes a ξ y a η los hemos caracterizado por los autovalores ξ' y η', o bien ξ^r y η' a los que pertenecen, junto con las letras c o d que también son necesarias. Igual como hacíamos hasta ahora con los autovectores de un único observable, de aquí en adelante siempre emplearemos los autovalores para designar a los autovectores comunes a varios observables como regla general.

El recíproco del teorema precedente dice que *si ξ y η son dos observables tales que los autoestados comunes a ambos constituyen un conjunto completo, entonces ξ y η conmutan.* Para demostrarlo tengamos en cuenta que si $|\xi'\eta'\rangle$ es un autoket común a ambos perteneciente a los autovalores ξ' y η', se tiene

$$(\xi\eta-\eta\xi)|\xi'\eta'\rangle = (\xi'\eta'-\eta'\xi')|\xi'\eta'\rangle = 0. \tag{49}$$

Pero como los autoestados comunes a ambos constituyen un conjunto completo, todo ket $|P\rangle$ puede ser expresado en función de autokets $|\xi'\eta'\rangle$ comunes a ξ y a η, para cada uno de los cuales se verifica (49), luego

$$(\xi\eta-\eta\xi)|P\rangle = 0$$

de modo que

$$\xi\eta-\eta\xi = 0.$$

La cualidad de los estados de ser autoestados comunes puede ser válida para más de dos observables, en cuyo caso tanto el primer teorema como su recíproco siguen siendo válidos, es decir, si cualquier conjunto de observables conmutan dos a dos, los autoestados comunes a todos ellos constituyen un conjunto completo y recíprocamente. Los argumentos empleados para demostrar los teoremas en el caso de dos observables pueden aplicarse a su vez para demostrar el caso general; por ejemplo, si tenemos tres observables que conmutan ξ, η y ζ, podemos expresar cualquier autoket común a ξ y η en función de autokets de ζ y ver que cada uno de esos autokets de ζ es a su vez autoket de ξ y de η. Con ello todo autoket común a ξ y a η queda expresado en función de autokets comunes a ξ, η y ζ, y puesto que todo ket puede expresarse en función de autokets comunes a ξ y η, resulta que todo ket se puede expresar en función de autokets comunes a ξ, η y ζ.

El teorema de ortogonalidad aplicado a este caso nos dice que cualquier par de autokets comunes a un conjunto de observables que conmutan, y que pertenezcan respectivamente a dos conjuntos de autovalores que difieran en algún elemento, son ortogonales.

Considerando el conjunto completo de los autoestados comunes a dos o más observables que conmutan, podemos elaborar una teoría de las funciones de ellos igual que hicimos en § 11 para las funciones de un único observable. Si ξ, η, ζ, ..., son observables que conmutan, definimos una función general de ellos f como el operador $f(\xi, \eta, \zeta, \ldots)$ que verifica

$$f(\xi, \eta, \zeta, \ldots)|\xi'\eta'\zeta'\ldots\rangle = f(\xi', \eta', \zeta', \ldots)|\xi'\eta'\zeta'\ldots\rangle, \tag{50}$$

para todo autoket $|\xi'\eta'\zeta'\ldots\rangle$ común a ξ, η, ζ, ..., y cuyos autovalores son ξ', η', ζ', ... En esta definición f es cualquier función con tal de que $f(a, b, c, \ldots)$ esté definida para todos los valores de a, b, c, ... que sean autovalores de ξ, η, ζ, ..., respectivamente. Del mismo modo que hicimos con las funciones de un único observable, definidas por (34), podemos demostrar que $f(\xi, \eta, \zeta, \ldots)$ está unívocamente determinado por (50) y que

$$\overline{f(\xi, \eta, \zeta, \ldots} = \bar{f}(\xi, \eta, \zeta, \ldots),$$

en correspondencia con (37). Asimismo, si $f(a, b, c, \ldots)$ es una función real, $f(\xi, \eta, \zeta, \ldots)$ es real y observable.

Podemos generalizar ahora los resultados (45) y (46). Dado un conjunto de observables ξ, η, ζ ..., que conmutan, podemos definir una función de ellos que sea igual a 1 para $\xi = a$, $\eta = b$, $\zeta = c$, ..., con a, b, c, ..., números reales cualesquiera, e igual a cero cuando no se verifique dicha condición. Esta función puede representarse por $\delta_{\xi a}$, $\delta_{\eta b}$, $\delta_{\zeta c}$, ..., ya que en realidad es igual al producto en cualquier orden de los factores $\delta_{\xi a}$, $\delta_{\eta b}$, $\delta_{\zeta c}$, ..., definidos cada uno de ellos como funciones de un solo observable, como puede comprobarse sustituyendo dicho producto en lugar de $f(\xi, \eta, \zeta, \ldots)$ en el primer miembro de (50). El valor esperado de dicha función en cualquier estado será igual a la probabilidad $P_{abc\ldots}$ de que ξ, η, ζ, ..., tengan respectivamente los valores a, b, c, ..., cuando el sistema esté en dicho estado. Por tanto, si dicho estado corresponde al ket normalizado $|x\rangle$, a partir de nuestra hipótesis general sobre la interpretación física, resulta

$$P_{abc\ldots} = \langle x|\delta_{\xi a}\delta_{\eta b}\delta_{\zeta c}\ldots|x\rangle. \tag{51}$$

$P_{abc\ldots}$ será nula a menos que cada uno de los números a, b, c, ..., sea autovalor del correspondiente observable. Si cada uno de los números a, b, c, ..., es un autovalor del correspondiente observable perteneciente a un dominio de autovalores de medida no nula, por regla general también será nula la probabilidad $P_{abc\ldots}$, pero en este caso podemos sustituir la exigencia de que el observable tenga exactamente un valor por la condición de que tenga un valor comprendido en un pequeño intervalo, lo que se consigue sustitu-

yendo los factores δ de (51) por un factor correspondiente a $\chi(\xi)$ de la ecuación (46). Llevando a cabo esta sustitución para cada uno de los observables ξ, η, ζ, ..., cuyos valores numéricos correspondientes a, b, c, ... formaban parte de un dominio de autovalores de medida no nula, obtendremos una probabilidad que en general ya no es nula.

Si varios observables conmutan, existen estados para los que todos los observables tienen valores particulares en el sentido indicado (al principio de la pág. 58), que son los autoestados comunes. Por tanto, *tiene sentido hablar de que varios observables que conmutan tienen un valor simultáneamente*. Aún más, vemos por (51) que para cualquier estado *tiene sentido hablar de la probabilidad de obtener resultados particulares al medir simultáneamente distintos observables que conmutan*. Esta nueva consecuencia es importante. En general no es posible llevar a cabo una observación de un sistema que esté en un estado definido sin alterar dicho estado y trastornarlo a fines de llevar a cabo una segunda observación. En el caso general, no tiene sentido hablar de que realizamos las dos mediciones simultáneamente. Sin embargo, la consecuencia anterior nos dice que, en el caso particular de que los dos observables conmuten, puede considerarse que las dos observaciones no interfieren entre sí, o que son *compatibles* en el sentido de que *podemos* considerar que las dos medidas se realizan simultáneamente, y hablar de la probabilidad de obtener resultados particulares cualesquiera en ellas. De hecho las dos observaciones pueden considerarse como una sola observación más complicada cuyo resultado viene dado por dos números en lugar de por uno. *Desde el punto de vista de la teoría general, cualquier par o más de observables que conmutan puede ser considerado como un único observable cuyo resultado de medida viene dado por dos o más números*. Los estados para los cuales dicha medición da con certeza un resultado particular son los autoestados comunes.

III

REPRESENTACIONES

14. Vectores básicos

En los capítulos anteriores hemos establecido un esquema algebraico con tres tipos de cantidades — los vectores bra, los vectores ket y los operadores lineales — y mediante dicho esquema hemos expresado algunas leyes fundamentales de la naturaleza. Se podría seguir desarrollando la teoría y aplicarla a problemas particulares considerando únicamente dichas cantidades abstractas. Sin embargo, para ciertos fines es preferible sustituir dichas cantidades por conjuntos de números con propiedades matemáticas análogas y hacer los desarrollos con ayuda de ellos. El procedimiento es similar al de elegir coordenadas en geometría, y tiene la ventaja de ser un método matemático más potente para resolver problemas particulares.

El procedimiento de sustituir las cantidades abstractas por números no es único. Igual que existen muchos sistemas de coordenadas en geometría, también hay muchas maneras de sustituir nuestras cantidades abstractas por números. A cada uno de estos procedimientos se le denomina *representación* y al conjunto de números que sustituyen a una cantidad abstracta le denominaremos el *representante* de dicha cantidad en la representación. Así pues el representante de una cantidad abstracta corresponde a las coordenadas de un objeto geométrico. Cuando estudiamos un problema particular en mecánica cuántica, podemos ahorrarnos trabajo empleando una representación en la que los representantes de las cantidades abstractas más importantes que intervienen en el problema sean lo más sencillas posible.

Para fijar una representación en el caso general, consideremos un conjunto completo de bras, o sea, un conjunto de bras tal que todo bra se pueda expresar linealmente en función de los bras del conjunto (bien como una suma de bras, bien como una integral, o como una integral más una suma). A estos bras les denominaremos *bras básicos* de la representación, y como veremos, serán suficientes para determinar completamente la representación.

Sea un ket $|a\rangle$ cualquiera y formemos los productos escalares con cada

representante de $|a\rangle$
$\langle\lambda_1|a\rangle, \langle\lambda_1|a\rangle, \cdots \langle\lambda_n|a\rangle$

uno de los bras básicos. Los números así obtenidos constituyen el representante de $|a\rangle$, y veamos que son suficientes para determinar completamente el ket $|a\rangle$. En efecto, si existiera otro ket $|a_1\rangle$ para el que dichos números fueran iguales, el producto escalar de la diferencia $|a\rangle - |a_1\rangle$ con cualquier bra básico sería nulo y, por tanto, también sería nulo el producto escalar con cualquier bra. Luego, $|a\rangle - |a_1\rangle$ sería nulo también.

Supongamos que los bras básicos están simbolizados por uno o más parámetros $\lambda_1, \lambda_2, \ldots, \lambda_u$, cada uno de los cuales puede tomar ciertos valores numéricos. Dicho bra básico lo escribiremos entonces $\langle\lambda_1\lambda_2\ldots\lambda_u|$ y el representante de $|a\rangle$ lo escribiremos $\langle\lambda_1\lambda_2\ldots\lambda_u|a\rangle$. Así pues, el representante está constituido por un conjunto de números en correspondencia con cada conjunto de valores de $\lambda_1, \lambda_2, \ldots, \lambda_u$, que forme parte de los respectivos campos de variabilidad. Un conjunto de números tal, constituye una *función* de las variables $\lambda_1, \lambda_2, \ldots, \lambda_u$. Por tanto, el representante de un ket puede considerarse bien como un conjunto de números o bien como una función de las variables que se emplean para especificar los bras básicos.

Si el número de estados independientes de nuestro sistema dinámico es finito e igual a n, será suficiente considerar n bras básicos que podremos designar con ayuda de un solo parámetro λ que tome los valores $1, 2, 3, \ldots, n$. En este caso, el representante de un ket $|a\rangle$ cualquiera está constituido por los n números $\langle 1|a\rangle, \langle 2|a\rangle, \langle 3|a\rangle, \ldots, \langle n|a\rangle$, que son precisamente las coordenadas del vector $|a\rangle$ referidas a un sistema de coordenadas ordinarias. Por tanto, el concepto de representante de un ket es una generalización del concepto de coordenadas de un vector ordinario, y cuando el número de dimensiones del espacio de los kets es finito, ambos coinciden.

En una representación general no es necesario que todos los bras básicos sean independientes. Pero en la mayoría de las representaciones que se utilizan en la práctica son independientes y además verifican la condición más severa de ser ortogonales dos a dos. En este caso se dice que la representación es una *representación ortogonal.*

Sea una representación ortogonal cuyos bras básicos $\langle\lambda_1\lambda_2\ldots\lambda_u|$ estén simbolizados por los parámetros $\lambda_1, \lambda_2, \ldots, \lambda_u$, que sólo pueden tomar valores reales. Dado un ket $|a\rangle$, podemos tomar su representante $\langle\lambda_1\lambda_2\ldots\lambda_u|a\rangle$, y considerar los números $\lambda_1\langle\lambda_1\lambda_2\ldots\lambda_u|a\rangle$, formados a partir de los anteriores, como el representante de un nuevo ket $|b\rangle$. La posibilidad de esta elección se deriva de que los números que forman el representante de un ket son independientes, dado que los bras básicos también lo son. Con ello el ket $|b\rangle$ queda definido por la ecuación

$$\langle\lambda_1\lambda_2\ldots\lambda_u|b\rangle = \lambda_1\langle\lambda_1\lambda_2\ldots\lambda_u|a\rangle.$$

Dicho ket $|b\rangle$ evidentemente es función lineal del ket $|a\rangle$ y, por lo tanto, puede ser considerado como el resultado de aplicar un operador lineal a $|a\rangle$. Si llamamos L_1 a dicho operador, resulta

$$|b\rangle = L_1\,|a\rangle$$

y en consecuencia $\langle\lambda_1\lambda_2\ldots\lambda_u|L_1|a\rangle = \lambda_1\langle\lambda_1\lambda_2\ldots\lambda_u|a\rangle$.

Esta ecuación es válida para todo ket $|a\rangle$, luego

$$\langle\lambda_1\lambda_2\ldots\lambda_u|L_1 = \lambda_1\langle\lambda_1\lambda_2\ldots\lambda_u|. \tag{1}$$

La ecuación (1) puede tomarse como la definición del operador lineal L_1. En ella vemos que *todo bra básico es un autobra de L_1 y el autovalor a que pertenece es el valor del parámetro λ_1.*

De la condición de que los bras básicos sean ortogonales se deduce que L_1 es real y que es un observable. En efecto sean $\lambda'_1, \lambda'_2, \ldots, \lambda'_u$ y $\lambda''_1, \lambda''_2, \ldots, \lambda''_u$ dos conjuntos de valores de los parámetros $\lambda_1, \lambda_2, \ldots, \lambda_u$. Poniendo los λ'_i en lugar de los λ_i en la ecuación (1) y multiplicando a la derecha por el conjugado imaginario del bra básico $\langle\lambda''_1\lambda''_2\ldots\lambda''_u|$, o sea por $|\lambda''_1\lambda''_2\ldots\lambda''_u\rangle$, resulta

$$\langle\lambda'_1\lambda'_2\ldots\lambda'_u|L_1|\lambda''_1\lambda''_2\ldots\lambda''_u\rangle = \lambda'_1\langle\lambda'_1\lambda'_2\ldots\lambda'_u|\lambda''_1\lambda''_2\ldots\lambda''_u\rangle.$$

Permutando los λ'_i y los λ''_i obtendremos

$$\langle\lambda''_1\lambda''_2\ldots\lambda''_u|L_1|\lambda'_1\lambda'_2\ldots\lambda'_u\rangle = \lambda''_1\langle\lambda''_1\lambda''_2\ldots\lambda''_u|\lambda'_1\lambda'_2\ldots\lambda'_u\rangle.$$

Dado que los bras básicos son ortogonales, los segundos miembros de estas dos ecuaciones se anulan excepto si $\lambda''_r = \lambda'_r$ para todo r desde 1 hasta u, en cuyo caso son iguales y además reales por serlo λ'_1. Luego, tanto si los λ''_r y los λ'_r son iguales como si no lo son, se tiene

$$\begin{aligned}\langle\lambda'_1\lambda'_2\ldots\lambda'_u|L_1|\lambda''_1\lambda''_2\ldots\lambda''_u\rangle &= \overline{\langle\lambda''_1\lambda''_2\ldots\lambda''_u|L_1|\lambda'_1\lambda'_2\ldots\lambda'_u\rangle}\\ &= \langle\lambda'_1\lambda'_2\ldots\lambda'_u|\bar{L}_1|\lambda''_1\lambda''_2\ldots\lambda''_u\rangle\end{aligned}$$

según la ecuación (4) de § 8. Puesto que los $\langle\lambda'_1\lambda'_2\ldots\lambda'_u|$ constituyen un conjunto completo de bras y los $|\lambda''_1\lambda''_2\ldots\lambda''_u\rangle$ un conjunto completo de kets, resulta que $L_1 = \bar{L}_1$. La otra condición requerida para que L_1 sea un observable, la de que sus autoestados constituyan un conjunto completo, evidentemente también se satisface, pues sus autobras son los bras básicos, que constituyen un conjunto completo.

Asimismo podríamos introducir otros operadores lineales $L_2, L_3, \ldots L_u$ multiplicando $\langle\lambda_1\lambda_2\ldots\lambda_u|a\rangle$ por los factores $\lambda_2, \lambda_3, \ldots \lambda_u$ respectivamente y considerando los conjuntos de números que así resultan como los representantes de nuevos kets. Análogamente podemos ver que los bras básicos son autobras de cada uno de los L_i, y dichos operadores son reales y observables. Los bras básicos son autobras comunes a todos los L_i, y dado que constituyen un conjunto completo, resulta, según el teorema de § 13, que los L_i conmutan dos a dos.

Vamos a demostrar ahora que *si $\xi_1, \xi_2, \ldots, \xi_u$ es un conjunto cualquiera de observables que conmutan, podemos formar con ellos una representación*

ortogonal en la que los bras básicos son autobras comunes a los $\xi_1, \xi_2, \ldots \xi_u$. Supongamos primero que no exista más que un solo autobra independiente común a $\xi_1, \xi_2, \ldots, \xi_u$ que pertenezca al conjunto de autovalores ξ'_1, ξ'_2, ξ'_u. En este caso tomaríamos como bras básicos los autobras comunes con factores numéricos arbitrarios. En virtud del teorema de ortogonalidad todos ellos serán ortogonales (cualquier par de ellos tendrá al menos algún autovalor distinto, y ello es suficiente para que sean ortogonales), y según el resultado de § 13 habrá suficientes para constituir un conjunto completo. Podemos designarlos adecuadamente mediante los autovalores $\xi'_1, \xi'_2, \ldots, \xi'_u$, a que pertenecen, y así escribiremos uno de ellos $\langle\xi'_1\xi'_2\ldots\xi'_u|$.

En el caso general en que para algunos conjuntos de autovalores existan varios autobras independientes comunes a $\xi_1, \xi_2, \ldots, \xi_u$, de todos los autobras comunes pertenecientes a los autovalores $\xi'_1\ \xi'_2, \ldots, \xi'_u$ tenemos que elegir solamente un subconjunto completo, de modo que cada bra sea ortogonal a los otros. (En este caso, la condición de completitud quiere decir que cualquier autobra común perteneciente a los autovalores $\xi'_1, \xi'_2, \ldots, \xi'_u$ se pueda expresar linealmente en función de los elementos del subconjunto.) Esto hemos de hacerlo para cada conjunto de autovalores $\xi'_1, \xi'_2, \ldots, \xi'_u$ y después reunir todos los elementos de cada subconjunto así obtenidos, y considerarlos como los bras básicos de la representación. Todos estos bras son ortogonales, pues si pertenecen a autovalores distintos son ortogonales en virtud del teorema de ortogonalidad, y los que pertenecen a un mismo conjunto de autovalores lo son por el modo como los hemos definido; además constituyen un conjunto completo de bras, ya que todo bra se puede expresar linealmente en función de autobras comunes y cada autobra común se puede expresar a su vez linealmente en función de los bras del subconjunto. Existen infinitas maneras de elegir dichos subconjuntos, y cada una de ellas constituye una representación ortogonal.

En este caso general para designar un bra básico, además de los autovalores $\xi'_1, \xi'_2, \ldots, \xi'_u$ a los que pertenece hemos de añadir ciertas variables reales $\lambda_1, \lambda_2, \ldots, \lambda_v$ para distinguir los vectores básicos que pertenecen al mismo conjunto de autovalores. Los bras básicos serán pues $\langle\xi'_1\xi'_2\ldots\xi'_u\lambda_1\lambda_2\ldots\lambda_v|$. Podemos ahora definir unos operadores lineales $L_1, L_2, \ldots, L_v$ relativos a las variables $\lambda_1, \lambda_2, \ldots, \lambda_v$ mediante ecuaciones análogas a la (1), y ver que los bras básicos son autobras de ellos, y que dichos operadores son reales y observables, que además conmutan dos a dos y con los ξ_i. Así resulta que los bras básicos son autobras comunes a todos los observables $\xi_1, \xi_2, \ldots, \xi_u, L_1, L_2, \ldots, L_v$ que conmutan dos a dos.

Definimos un *conjunto completo de observables que conmutan* como un conjunto de observables que conmutan dos a dos, para el que no existe más que un único autoestado común perteneciente a un conjunto cualquiera

de autovalores. Con esta definición los observables $\xi_1, \xi_2, \ldots, \xi_u, L_1, L_2, \ldots, L_v$ constituyen un conjunto completo de observables, ya que no existe más que un solo autobra independiente perteneciente a los autovalores $\xi'_1, \xi'_2, \ldots, \xi'_u, \lambda_1, \lambda_2, \ldots, \lambda_v$, que es precisamente el bra básico correspondiente. Y lo mismo ocurre con los operadores $L_1, L_2, \ldots, L_v$ definidos por la ecuación (1) y las consideraciones que la siguen. Con ayuda de esta definición, todos estos resultados se pueden formular brevemente así:

(i) Los bras básicos de una representación ortogonal son autobras comunes a un conjunto completo de observables que conmutan.

(ii) Dado un conjunto completo de observables que conmutan, podemos construir una representación ortogonal en la que los bras básicos son todos autobras comunes a todos los observables del conjunto completo.

(iii) Dado un conjunto cualquiera de observables que conmutan, siempre es posible añadirle algunos observables más para formar un conjunto completo.

(iv) Para designar los bras básicos de una representación ortogonal, es conveniente utilizar los autovalores de los observables del conjunto completo del que dichos bras son autobras comunes.

A los imaginarios conjugados de los bras básicos de una representación los denominaremos *kets básicos* de la representación. Así, si los bras básicos son $\langle\lambda_1\lambda_2\ldots\lambda_u|$, los kets básicos serán $|\lambda_1\lambda_2\ldots\lambda_u\rangle$. El representante de un bra $\langle b|$ vendrá dado por su producto escalar con cada uno de los kets fundamentales, es decir, por $\langle b|\lambda_1\lambda_2\ldots\lambda_u\rangle$. Igual como ocurría con el representante de un ket, también a éste se le puede considerar como un conjunto de números o como una función de las variables $\lambda_1, \lambda_2, \ldots \lambda_u$. Se tiene

$$\langle b|\lambda_1\lambda_2\ldots\lambda_u\rangle = \overline{\langle\lambda_1\lambda_2\ldots\lambda_u|b\rangle},$$

lo que nos dice que *el representante de un bra es igual al complejo conjugado del representante del ket imaginario conjugado.* En una representación ortogonal en que los bras básicos son autobras comunes a un conjunto completo de observables que conmutan $\xi_1, \xi_2, \ldots, \xi_u$, los kets básicos serán a su vez autokets comunes a $\xi_1, \xi_2, \ldots, \xi_u$.

Hasta ahora no hemos dicho nada acerca de las longitudes de los kets básicos. En una representación ortogonal es más natural normalizar los kets básicos que no tomarlos con longitudes arbitrarias, y con ello introducimos una simplificación más en la representación. Pero sólo será posible normalizarlos cuando todos los parámetros que se necesitan para designarlos tomen únicamente valores discretos. Si alguno de los parámetros indicados es una variable continua que puede tomar todos los valores de un cierto dominio de medida no nula, entonces los vectores básicos son autovectores de un cierto observable que pertenecen a autovalores de un conjunto de medida no nula y, por tanto, según la discusión de § 10 (véase pág. 51) son de

longitud infinita. En este caso es necesario otro procedimiento para fijar los factores numéricos por los que pueden multiplicarse los vectores básicos. Para estudiar de un modo preciso esta cuestión es necesaria una nueva notación matemática, que se da en la sección siguiente.

15. *La función* δ

Los razonamientos de § 10 nos llevaron a considerar cantidades que implican un tipo de infinito especial. Para manejar dichos infinitos con una notación rigurosa introducimos la cantidad $\delta(x)$ que depende de un parámetro x y satisface las condiciones

$$\left.\begin{aligned} &\int_{-\infty}^{\infty} \delta(x)\, dx = 1 \\ &\delta(x) = 0 \text{ para } x \neq 0. \end{aligned}\right\} \tag{2}$$

Si queremos tener una imagen de $\delta(x)$, consideremos una función de la variable real x que sea nula fuera de un pequeño dominio de amplitud ϵ alrededor del origen $x = 0$ y que en el interior de este dominio sea igual a uno. No importa la forma exacta de la función en el interior de este dominio, con tal de que no sufra en él variaciones innecesariamente bruscas (por ejemplo, con tal de que la función sea siempre del orden de ϵ^{-1}). Pasando al límite para $\epsilon \to 0$, esta función tenderá a confundirse con $\delta(x)$.

$\delta(x)$ no es una función de x según la definición matemática ordinaria de función — que le exigiría tener un valor definido para cada punto de su dominio — sino algo más general que llamaremos 'función impropia' * para destacar su diferencia con las funciones definidas de modo ordinario. Por tanto, $\delta(x)$ no es una cantidad que pueda usarse en análisis matemático con tanta generalidad como las funciones ordinarias, y su uso debe restringirse a ciertos tipos de expresiones sencillas para las que sea evidente que no puede dar lugar a inconsecuencias lógicas.

La propiedad más importante de $\delta(x)$ puede expresarse con la siguiente ecuación,

$$\int_{-\infty}^{\infty} f(x)\delta(x)\, dx = f(0), \tag{3}$$

en la que $f(x)$ es cualquier función continua de x. Es fácil ver la validez de esta ecuación a partir de la imagen de $\delta(x)$ que acabamos de dar. El primer miembro de (3) sólo puede depender de los valores $f(x)$ muy próximos al

* La $\delta(x)$ que introdujo el autor de este modo heurístico ha dado lugar a la teoría de distribuciones de L. Schwarz, dentro de la cual queda definida con todo rigor. En honor al autor, $\delta(x)$ se denomina δ de Dirac. *(N. del T.)*

origen, de forma que sin error notable podemos sustituir $f(x)$ por su valor en el origen $f(0)$. La ecuación (3) resulta entonces como consecuencia de la primera de las ecuaciones (2). Haciendo un cambio de origen en (3), podemos deducir la fórmula

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)\ dx = f(a), \tag{4}$$

en la que a es un número real cualquiera. Luego, *el resultado de multiplicar una función de x por $\delta(x-a)$ e integrar para todo x es equivalente a sustituir x por a.* Este resultado general es válido aunque la función de x no sea una función numérica, sino un vector o un operador lineal dependientes de x.

El intervalo de integración en (3) y (4) no ha de ser necesariamente de $-\infty$ a $+\infty$, sino que puede ser cualquier dominio que incluya el punto crítico en el que la función δ no se anula. De ahora en adelante omitiremos de ordinario los límites de integración en ese tipo de ecuaciones, y daremos por sobreentendido que las integrales se extienden a un dominio adecuado.

Las ecuaciones (3) y (4) muestran que aunque una función impropia considerada aisladamente no tenga un valor bien definido, cuando aparece como factor en un integrando, la integral sí que tiene un valor bien definido. En la teoría cuántica, siempre que aparece una función impropia aparece dentro de una expresión que finalmente colòcaremos en un integrando. Por consiguiente, sería posible volver a escribir la teoría de forma que las funciones impropias aparecieran únicamente en integrandos. Y así se podría prescindir completamente de las funciones impropias. Así pues, el uso de las funciones impropias no implica ninguna falta de rigor en nuestra teoría, y no es más que una notación cómoda que nos permite expresar en forma concisa ciertas relaciones que, si fuera necesario, podríamos también escribir sin emplear funciones impropias, aunque esto resultaría muy engorroso y fácilmente oscurecería la idea central.

Podemos definir la función δ todavía de otro modo, como la derivada $\varepsilon'(x)$ de la función $\varepsilon(x)$ definida así:

$$\left.\begin{aligned} \varepsilon(x) &= 0 \quad (x < 0) \\ &= 1 \quad (x > 0). \end{aligned}\right\} \tag{5}$$

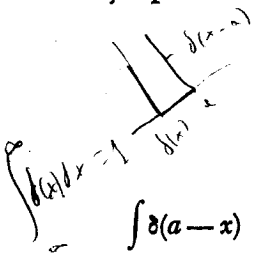
Podemos comprobar que esta nueva definición es equivalente a la anterior, sustituyendo $\delta(x)$ por $\varepsilon'(x)$ en el primer miembro de (3) e integrando por partes. Si g_1 y g_2 son dos números positivos cualesquiera, resulta

$$\begin{aligned} \int_{-g_2}^{g_1} f(x)\varepsilon'(x)\ dx &= [f(x)\varepsilon(x)]_{-g2}^{g1} - \int_{-g_2}^{g_1} f'(x)\varepsilon(x)\ dx \\ &= f(g_1) - \int_0^{g_1} f'(x)\ dx \\ &= f(0), \end{aligned}$$

de acuerdo con (3). La función δ aparece al derivar una función discontinua.

Con las funciones δ se pueden escribir algunas ecuaciones elementales. Tales ecuaciones son esencialmente reglas de manipulación para desarrollos algebraicos en que intervengan funciones δ. El sentido de cualquiera de estas ecuaciones es el de que sus dos miembros dan el mismo resultado empleados como factores en un integrando.

Ejemplos de estas ecuaciones son:



$$\delta(-x) = \delta(x) \tag{6}$$

$$x\delta(x) = 0, \tag{7}$$

$$\delta(ax) = a^{-1}\delta(x) \quad (a > 0), \tag{8}$$

$$\delta(x^2 - a^2) = \tfrac{1}{2}a^{-1}\{\delta(x-a) + \delta(x+a)\} \quad (a > 0), \tag{9}$$

$$\int \delta(a-x)\ dx\ \delta(x-b) = \delta(a-b), \tag{10}$$

$$f(x)\delta(x-a) = f(a)\delta(x-a). \tag{11}$$

La ecuación (6), que enuncia simplemente el hecho de que $\delta(x)$ es una función par de la variable x, es trivial. Para comprobar (7) consideremos una función de x continua $f(x)$ cualquiera. Para ella

$$\int f(x)x\delta(x)\ dx = 0,$$

según (3). Luego, $x\delta(x)$ como factor en un integrando es equivalente a cero, y ese es precisamente el significado de (7). (8) y (9) pueden comprobarse mediante razonamientos elementales análogos. Para comprobar (10) consideremos una función de a continua $f(a)$ cualquiera. Para ella

$$\int f(a)\ da \int \delta(a-x)\ dx\ \delta(x-b) = \int \delta(x-b)\ dx \int f(a)\ da\ \delta(a-x).$$
$$= \int \delta(x-b)\ dx\ f(x) = \int f(a)\ da\ \delta(a-b).$$

Luego, los dos miembros de (10) son equivalentes cuando se emplean como factores de un integrando con a como variable de integración. Igualmente se puede demostrar que también son equivalentes como factores en un integrando con b como variable de integración, de forma que la ecuación (10) queda comprobada desde ambos puntos de vista. La ecuación (11) también puede comprobarse fácilmente desde ambos puntos de vista con ayuda de (4).

La ecuación (10) resulta también como aplicación de (4) tomando $f(x) = \delta(x-b)$. He aquí una ilustración del hecho de que muchas veces podamos usar una función impropia como si fuera una función continua ordinaria sin que ello nos lleve a ningún resultado incorrecto.

La ecuación (7) nos indica que al dividir ambos miembros de una ecuación por una variable x que pueda tomar el valor cero, debemos siempre

añadir a uno de los miembros un múltiplo arbitrario de $\delta(x)$, es decir, que de la ecuación

$$A = B, \tag{12}$$

no se puede deducir

$$A/x = B/x,$$

sino sólo

$$A/x = B/x + c\delta(x), \tag{13}$$

siendo c un número desconocido.

Como aplicación ilustrativa de la función $\delta(x)$, consideremos la derivada de $\log x$. La forma usual

$$\frac{d}{dx} \log x = \frac{1}{x} \tag{14}$$

debe ser examinada en el entorno de $x = 0$. Para que la función inversa $1/x$ quede definida en el entorno de $x = 0$ (en el sentido de una función impropia) tenemos que imponerle una condición supletoria; por ejemplo que se anule su integral de $-\varepsilon$ a ε. Con esta condición supletoria, la integral del segundo miembro de (14) de $-\varepsilon$ a ε se anula, mientras que la del primer miembro resulta ser igual a $\log(-1)$, de donde (14) no es una ecuación correcta. Para corregirla debemos recordar que $\log x$ tiene un término imaginario puro para los valores negativos de x, que si tomamos valores principales es igual a $i\pi$. Al pasar x por el valor cero, el término imaginario puro se anula de modo discontinuo. La derivada de dicho término resulta igual a $-i\pi\delta(x)$, de forma que en vez de (14) habrá que escribir

$$\frac{d}{dx} \log x = \frac{1}{x} - i\pi\delta(x). \tag{15}$$

Esta combinación especial de función inversa y función δ que aparece en (15) juega un papel importante en la teoría cuántica de colisiones (véase § 50).

16. *Propiedades de los vectores básicos*

Empleando la notación de la función δ, podemos continuar nuestra teoría de las representaciones. Supongamos en primer lugar que un único observable constituye por sí solo un conjunto completo de observables que conmutan — para que esto se cumpla basta que ξ no tenga más que un solo autoestado para cada autovalor ξ' —, y construyamos una representación ortogonal cuyos vectores básicos sean los autovectores de ξ que escribiremos $\langle\xi'|$, $|\xi'\rangle$.

Si los autovalores de ξ son discretos, podemos normalizar los vectores básicos, y así tendremos

$$\langle \xi' | \xi'' \rangle = 0 \quad (\xi' \neq \xi''),$$
$$\langle \xi' | \xi' \rangle = 1.$$

Estas dos ecuaciones se pueden combinar en una única ecuación

$$\langle \xi' | \xi'' \rangle = \delta_{\xi'\xi''}, \tag{16}$$

en la que el símbolo δ con dos subíndices, que en adelante utilizaremos con frecuencia, vale

$$\left.\begin{aligned} \delta_{rs} &= 0 \quad \text{si} \quad r \neq s \\ &= 1 \quad \text{si} \quad r = s. \end{aligned}\right\} \tag{17}$$

Si los autovalores de ξ son continuos, no podemos normalizar los vectores básicos. Pero si consideramos la cantidad $\langle \xi' | \xi'' \rangle$ con ξ' fija y ξ'' variable, y recordamos nuestras reflexiones a propósito de la expresión (29) de § 10, vemos que dicha cantidad se anula para $\xi'' \neq \xi'$ y que su integral a lo largo de un intervalo de ξ'' que contenga a ξ' toma un valor finito que llamaremos c. Tendremos, por tanto,

$$\langle \xi' | \xi'' \rangle = c\delta(\xi' - \xi'').$$

Según (30) de § 10, c es un número positivo. Puede variar con ξ', por lo que conviene llamarle $c(\xi')$ o c' para abreviar. Obtenemos pues la ecuación

$$\langle \xi' | \xi'' \rangle = c'\delta(\xi' - \xi''). \tag{18}$$

Hubiéramos podido obtener también

$$\langle \xi' | \xi'' \rangle = c''\delta(\xi' - \xi''), \tag{19}$$

en la que c'' es una abreviación de $c(\xi'')$. Según (11) los segundos miembros de (18) y (19) son iguales.

Pasemos a otra representación cuyos vectores básicos sean también autovectores de ξ, por lo que los nuevos vectores básicos serán múltiplos de los anteriores. Llamando a los nuevos vectores básicos $\langle \xi'^* |$, $| \xi'^* \rangle$, con el índice supletorio * para distinguirlos de los anteriores, tendremos

$$\langle \xi'^* | = k' \langle \xi' |, \qquad | \xi'^* \rangle = \overline{k'} | \xi' \rangle,$$

siendo k' — abreviación de $k(\xi')$ — un cierto número que depende de ξ'. Calculemos

$$\langle \xi'^* | \xi''^* \rangle = k'\overline{k''} \langle \xi' | \xi'' \rangle = k'\overline{k''} c'\delta(\xi' - \xi'')$$

aplicando (18). Según (11) podemos escribirlo

$$\langle \xi'^* | \xi''^* \rangle = k'\overline{k'} c'\delta(\xi' - \xi'')$$

Si elegimos k' de forma que su módulo sea $c'^{-\frac{1}{2}}$ — siempre lo podremos hacer ya que c' es positivo — lograremos que sea

$$\langle \xi'^* | \xi''^* \rangle = \delta(\xi' - \xi''). \tag{20}$$

De esta manera hemos fijado las longitudes de los nuevos vectores básicos para hacer la representación lo más simple posible. El proceso por el que hemos fijado las longitudes de estos vectores es en cierta manera análogo a la normalización de los vectores básicos en el caso en que los ξ' eran discretos, y la ecuación (20) es de la misma forma que la (16) pero con la función $\delta(\xi' - \xi'')$ en lugar del símbolo $\delta_{\xi'\xi''}$. En adelante manejaremos la nueva representación, y suprimiremos los índices * para simplificar la escritura. Así pues la ecuación (20) la escribiremos sencillamente

$$\langle \xi' | \xi'' \rangle = \delta(\xi' - \xi''). \tag{21}$$

Podemos desarrollar nuestra teoría con un paralelismo perfecto para ambos casos discreto y continuo. Para el caso discreto obtendremos, aplicando (16),

$$\sum_{\xi'} |\xi'\rangle\langle\xi'|\xi''\rangle = \sum_{\xi'} |\xi'\rangle \delta_{\xi'\xi''} = |\xi''\rangle,$$

donde los sumatorios se extienden a todos los autovalores. Esta ecuación es válida para cualquier ket básico $|\xi''\rangle$ y, por consiguiente, como los kets básicos forman un conjunto completo, tendremos

$$\sum_{\xi'} |\xi'\rangle\langle\xi'| = 1. \tag{22}$$

Esta es una ecuación muy útil, que expresa una propiedad importante de los vectores básicos: *si $|\xi'\rangle$ se multiplica a la derecha por $\langle\xi'|$ resulta un operador lineal que sumado para todo ξ' equivale al operador unidad.* Las ecuaciones (16) y (22) nos dan las propiedades fundamentales de los vectores básicos para el caso discreto.

Análogamente para el caso continuo obtendremos, aplicando (21),

$$\int |\xi'\rangle \, d\xi' \, \langle\xi'|\xi''\rangle = \int |\xi'\rangle \, d\xi' \delta(\xi' - \xi'') = |\xi''\rangle \tag{23}$$

en la que hemos hecho uso de (4) con un vector ket en lugar de la función $f(x)$, y donde la integral está extendida a todo el intervalo de los autovalores. Al ser esto válido para cualquier ket básico $|\xi''\rangle$, será

$$\int |\xi'\rangle \, d\xi' \, \langle\xi'| = 1. \tag{24}$$

Esta ecuación es de la misma forma que (22) pero con una integral en lugar de una suma. Las ecuaciones (21) y (24) nos dan las propiedades fundamentales de los vectores básicos para el caso continuo.

Las ecuaciones (22) y (24) permiten desarrollar cualquier bra o ket en

función de los vectores básicos. Por ejemplo, para el ket $|P\rangle$ en el caso discreto, multiplicando (22) a la derecha por $|P\rangle$, obtendremos

$$|P\rangle = \sum_{\xi'} |\xi'\rangle\langle\xi'|P\rangle, \tag{25}$$

que nos da $|P\rangle$ expresado en función de los $|\xi'\rangle$ y nos muestra que los coeficientes de ese desarrollo $\langle\xi'|P\rangle$ son precisamente los números que constituyen el representante de $|P\rangle$. Análogamente, para el caso continuo

$$|P\rangle = \int |\xi'\rangle\, d\xi'\, \langle\xi'|P\rangle, \tag{26}$$

que nos da $|P\rangle$ como una integral extendida a los $|\xi'\rangle$ con coeficientes en el integrando que constituyen también el representante $\langle\xi'|P\rangle$ de $|P\rangle$. Las ecuaciones conjugadas imaginarias de (25) y (26) nos darán el vector bra $\langle P|$ expresado en función de los bras básicos.

Los métodos matemáticos que manejamos ahora nos permiten desarrollar cualquier ket en el caso continuo como una integral de autokets de ξ. Cuando no empleábamos la notación de la función δ, el desarrollo general de un ket consistía en una integral más una suma, tal como la ecuación (25) de § 10. Pero la función δ nos permite ahora sustituir la suma por una integral, cuyo integrando consta de una serie de términos, cada uno de los cuales contiene una función δ como factor. Por ejemplo, el autoket $|\xi''\rangle$ puede sustituirse por una integral de autokets según la segunda de las ecuaciones (23).

Si $\langle Q|$ es un bra cualquiera y $|P\rangle$ un ket cualquiera, aplicando de nuevo (22) y (24) obtendremos, para ξ' discretos

$$\langle Q|P\rangle = \sum_{\xi'} \langle Q|\xi'\rangle\langle\xi'|P\rangle \tag{27}$$

y para ξ' continuos

$$\langle Q|P\rangle = \int \langle Q|\xi'\rangle\, d\xi'\, \langle\xi'|P\rangle \tag{28}$$

Estas ecuaciones expresan el producto escalar de $\langle Q|$ y $|P\rangle$ en función de sus representantes $\langle Q|\xi'\rangle$ y $\langle\xi'|P\rangle$. La ecuación (27) es exactamente la fórmula usual del producto escalar de dos vectores en función de sus coordenadas, y (28) es la modificación natural de esta fórmula para el caso de ξ' continuo, con una integral en vez de un sumatorio.

La generalización de las consideraciones precedentes para el caso en que ξ tenga a la vez autovalores discretos y continuos no ofrece ninguna dificultad. Usando ξ^r y ξ^s para indicar autovalores discretos, y ξ' y ξ'' para indicar autovalores continuos, obtendremos el conjunto de ecuaciones

$$\langle\xi^r|\xi^s\rangle = \delta_{\xi^r\xi^s}, \qquad \langle\xi^r|\xi'\rangle = 0, \qquad \langle\xi'|\xi''\rangle = \delta(\xi'-\xi'') \tag{29}$$

como generalización de (16) o (21). Estas ecuaciones expresan que los vectores básicos son todos ortogonales, y que los correspondientes a autovalores

continuos tienen fijadas sus longitudes según el mismo criterio con que dedujimos (20). A partir de (29) podemos obtener la generalización de (22) o (24),

$$\sum_{\xi^r} |\xi^r\rangle\langle\xi^r| + \int |\xi'\rangle\, d\xi'\, \langle\xi'| = 1, \tag{30}$$

cuya integral está extendida a todo el dominio de los autovalores continuos. De (30) resulta inmediatamente

$$|P\rangle = \sum_{\xi^r} |\xi^r\rangle\langle\xi^r|P\rangle + \int |\xi'\rangle\, d\xi'\, \langle\xi'|P\rangle \tag{31}$$

que es la generalización de (25) o (26), y

$$\langle Q|P\rangle = \sum_{\xi^r} \langle Q|\xi^r\rangle\langle\xi^r|P\rangle + \int \langle Q|\xi'\rangle\, d\xi'\, \langle\xi'|P\rangle \tag{32}$$

que generaliza a (27) o (28).

Pasemos al caso general en que haya varios observables que conmutan $\xi_1, \xi_2, \ldots, \xi_u$ que forman un conjunto completo, y construyamos una representación ortogonal cuyos vectores básicos, que escribiremos $\langle\xi'_1\ldots\xi'_u|$, $|\xi'_1\ldots\xi'_u\rangle$, sean autovectores comunes a todos ellos. Supongamos que $\xi_1, \xi_2, \ldots, \xi_v$ $(v \leqslant u)$ tengan autovalores discretos y $\xi_{v+1}, \ldots, \xi_u$ tengan autovalores continuos.

Consideremos el producto escalar $\langle\xi'_1\ldots\xi'_v\xi'_{v+1}\ldots\xi'_u|\xi'_1\ldots\xi'_v\xi'_{v+1}\xi''_u\rangle$. Según el teorema de ortogonalidad debe ser nulo a no ser que $\xi''_s = \xi'_s$ para cada $s = v+1, \ldots, u$. Si generalizamos las consideraciones que hicimos sobre la expresión (29) de § 10 para aplicarlas a autovectores comunes a un conjunto de varios observables que conmutan y generalizando también el axioma (30), la integral $(u-v)$-ésima de dicho producto escalar respecto a cada ξ''_s a lo largo de un intervalo que contenga el valor ξ'_s es un número finito positivo. Llamando c' a dicho número —el acento indica que c' es función de $\xi'_1, \ldots, \xi'_v, \xi'_{v+1}, \ldots, \xi'_u$—, nuestros resultados pueden expresarse mediante la ecuación

$$\langle\xi'_1\ldots\xi'_v\xi'_{v+1}\ldots\xi'_u|\xi'_1\ldots\xi'_v\xi''_{v+1}\ldots\xi''_u\rangle = c'\,\delta(\xi'_{v+1}-\xi''_{v+1})\ldots\delta(\xi'_u-\xi''_u), \tag{33}$$

en la que figura un factor δ en el segundo miembro para cada valor de s desde $v+1$ a u. Podemos cambiar ahora las longitudes de nuestros vectores básicos, mediante un procedimiento análogo al que nos condujo a (20), de forma que resulte c' igual a uno. Usando de nuevo el teorema de ortogonalidad, obtenemos por fin

$$\langle\xi'_1\ldots\xi'_u|\xi''_1\ldots\xi''_u\rangle = \delta_{\xi'_1\xi''_1}\ldots\delta_{\xi'_v\xi''_v}\delta(\xi'_{v+1}-\xi''_{v+1})\ldots\delta(\xi'_u-\xi''_u), \tag{34}$$

en cuyo segundo miembro figura un símbolo δ con dos subíndices para

cada ξ de autovalores discretos, y una función δ para cada ξ de autovalores continuos. Esta ecuación generaliza la (16) o (21) para el caso de un conjunto completo de varios observables que conmutan.

De (34) se puede deducir la generalización de (22) o (24)

$$\sum_{\xi'_1 \cdots \xi'_v} \int \cdots \int |\xi'_1 \ldots \xi'_u\rangle \, d\xi'_{v+1} \ldots d\xi'_u \, \langle \xi'_1 \ldots \xi'_u| = 1, \tag{35}$$

en la que figura una integral $(u - v)$-ésima extendida a todas las ξ' de autovalores continuos y un sumatorio para todas las ξ' de autovalores discretos. Las ecuaciones (34) y (35) nos dan las propiedades fundamentales de los vectores básicos para el caso que estamos estudiando. A partir de (35) podemos escribir inmediatamente la generalización de (25) o (26) y la de (27) o (28).

El caso que acabamos de considerar se puede generalizar todavía, permitiendo que algunos de los observables ξ tengan a la vez autovalores discretos y continuos. Las modificaciones que hay que introducir en las ecuaciones son inmediatas, pero no las introduciremos porque son bastante engorrosas de escribir en forma general.

Hay problemas para los que en lugar de hacer que la c' de la ecuación (33) valga la unidad, resulta más conveniente hacerla igual a una cierta función de las ξ'. Si llamamos ρ'^{-1} a dicha función, en lugar de (34) tendremos

$$\langle \xi'_1 \ldots \xi'_u | \xi''_1 \ldots \xi''_u \rangle = \rho'^{-1} \delta_{\xi'_1 \xi''_1} \cdots \delta_{\xi'_v \xi''_v} \delta(\xi'_{v+1} - \xi''_{v+1}) \ldots \delta(\xi'_u - \xi''_u), \tag{36}$$

y en lugar de (35) tendremos

$$\sum_{\xi'_1 \cdots \xi'_v} \int \cdots \int |\xi'_1 \ldots \xi'_u\rangle \, \rho' \, d\xi'_{v+1} \ldots d\xi'_u \langle \xi'_1 \ldots \xi'_u| = 1. \tag{37}$$

ρ' suele llamarse *función de peso* de la representación, por ser $\rho' \, d\xi'_{v+1} \ldots d\xi'_u$ el 'peso' que se atribuye a un pequeño elemento de volumen del espacio de variables $\xi'_{v+1}, \ldots, \xi'_u$.

Todas las representaciones que hemos considerado anteriormente tienen la función de peso unidad. La introducción de una función de peso no unidad es enteramente cuestión de conveniencia, pues la representación no resulta con ello más potente desde el punto de vista matemático. Los bras básicos $\langle \xi'_1 \ldots \xi'^*_u|$ de una representación de función de peso ρ' se relacionan con los bras básicos $\langle \xi'_1 \ldots \xi'_u|$ de la representación correspondiente de función de peso unidad mediante

$$\langle \xi'_1 \ldots \xi'^*_u| = \rho'^{-\frac{1}{2}} \langle \xi'_1 \ldots \xi'_u|, \tag{38}$$

como es fácil comprobar. Un ejemplo de representación muy útil con función de peso no unidad es cuando dos ξ' son los ángulos polar y azimu-

tal θ y ϕ que dan la dirección en un espacio tridimensional. Conviene tomar entonces $\rho' = \text{sen}\ \theta'$, con lo que en (37) aparecerá el elemento de ángulo sólido sen $\theta'\ d\theta' d\phi'$.

17. *Representación de operadores lineales*

Vimos en § 14 cómo se pueden representar los vectores ket y bra mediante un conjunto de números. Ahora hemos de hacer lo mismo con los operadores lineales, para tener así un esquema completo que nos permita representar todas nuestras cantidades abstractas mediante conjuntos de números. Para ello podemos utilizar los mismos vectores básicos de § 14.

Supongamos que los vectores básicos son autovectores comunes a un conjunto completo de observables que conmutan $\xi_1, \xi_2, \ldots, \xi_u$. Dado un operador lineal α cualquiera, tomemos un bra básico general $\langle \xi'_1 \ldots \xi'_u|$ y un ket básico general $|\xi''_1 \ldots \xi''_u\rangle$, y formemos los números

$$\langle \xi'_1 \ldots \xi'_u | \alpha | \xi''_1 \ldots \xi''_u \rangle. \tag{39}$$

Estos números son suficientes para determinar α completamente, puesto que en primer lugar determinan el ket transformado $\alpha|\xi''_1 \ldots \xi''_u\rangle$ (ya que dan su representante), y al determinar el ket transformado para cada uno de los kets básicos $|\xi''_1 \ldots \xi''_u\rangle$ queda determinado α. Los números (39) constituyen el *representante* de un operador lineal α o de una variable dinámica α. Este representante es más complicado que el de un vector ket o bra, ya que encierra los parámetros que designan a dos vectores básicos en lugar de uno.

Examinemos la forma de estos números en casos sencillos. Consideremos el caso en que un solo observable ξ constituye por sí solo un conjunto completo, y supongamos que sus autovalores ξ' sean discretos. Entonces el representante de α está constituido por el conjunto discreto de números $\langle \xi' | \alpha | \xi'' \rangle$. Al escribir explícitamente estos números, es natural distribuirlos en una disposición bidimensional así:

$$\begin{pmatrix} \langle \xi^1|\alpha|\xi^1\rangle & \langle \xi^1|\alpha|\xi^2\rangle & \langle \xi^1|\alpha|\xi^3\rangle & \cdot & \cdot \\ \langle \xi^2|\alpha|\xi^1\rangle & \langle \xi^2|\alpha|\xi^2\rangle & \langle \xi^2|\alpha|\xi^3\rangle & \cdot & \cdot \\ \langle \xi^3|\alpha|\xi^1\rangle & \langle \xi^3|\alpha|\xi^2\rangle & \langle \xi^3|\alpha|\xi^3\rangle & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{40}$$

siendo $\xi^1, \xi^2, \xi^3, \ldots$ los autovalores de ξ. Este tipo de disposición se denomina *matriz*, y sus números, *elementos* de la matriz. Hacemos la convención de que los elementos que figuren en una misma fila pertenezcan siempre al mismo bra básico, y los de una misma columna a un mismo ket básico.

Un elemento $\langle \xi' | \alpha | \xi' \rangle$ correspondiente a dos vectores básicos con el

mismo índice se llama *elemento diagonal* de la matriz, pues todos los elementos de ese tipo están sobre una diagonal. Si hacemos α igual a la unidad, según (16), todos los elementos diagonales valdrán la unidad y todos los demás elementos serán nulos. Dicha matriz recibe el nombre de *matriz unidad.*

Si α es real, tenemos

$$\langle \xi' | \alpha | \xi'' \rangle = \overline{\langle \xi'' | \alpha | \xi' \rangle}. \tag{41}$$

Estas condiciones exigen que todos los elementos diagonales de la matriz (40) sean reales y que cada uno de los otros sea igual al complejo conjugado de su simétrico respecto a la diagonal. Una matriz de este tipo se denomina *matriz hermítica.*

Si hacemos α igual a ξ, obtenemos como elemento general de la matriz

$$\langle \xi' | \xi | \xi'' \rangle = \xi' \langle \xi' | \xi'' \rangle = \xi' \delta_{\xi' \xi''} \tag{42}$$

O sea, que todos los elementos que no están en la diagonal son nulos. Una matriz así se denomina *matriz diagonal.* Sus elementos diagonales son precisamente los autovalores de ξ. Generalizando, si hacemos α igual a $f(\xi)$ — una función cualquiera de ξ — obtenemos

$$\langle \xi' | f(\xi) | \xi'' \rangle = f(\xi') \delta_{\xi' \xi''}, \tag{43}$$

y la matriz que se obtiene también es diagonal.

Tratemos de hallar el representante del producto $\alpha\beta$ de los operadores lineales α y β en función de los representantes de los factores. Usando la ecuación (22) con ξ''' en vez de ξ' obtenemos

$$\begin{aligned} \langle \xi' | \alpha\beta | \xi'' \rangle &= \langle \xi' | \alpha \sum_{\xi'''} | \xi''' \rangle \langle \xi''' | \beta | \xi'' \rangle \\ &= \sum_{\xi'''} \langle \xi' | \alpha | \xi''' \rangle \langle \xi''' | \beta | \xi'' \rangle, \end{aligned} \tag{44}$$

que nos da el resultado que buscábamos. La ecuación (44) expresa que la matriz formada por los elementos $\langle \xi' | \alpha\beta | \xi'' \rangle$ es igual al producto de las matrices formadas respectivamente por los elementos $\langle \xi' | \alpha | \xi'' \rangle$ y $\langle \xi' | \beta | \xi'' \rangle$, obtenido por la regla matemática ordinaria de multiplicación de matrices. Esta regla nos da como elemento perteneciente a la fila r-ésima y a la columna s-ésima de la matriz producto, la suma de los productos de cada elemento de la fila r-ésima de la primera matriz por el correspondiente elemento de la columna s-ésima de la segunda matriz. La multiplicación de matrices, igual que la de operadores lineales no es conmutativa.

Para este primer caso de un solo observable ξ con autovalores discretos podemos resumir nuestros resultados del modo siguiente:

(i) *Todo operador lineal está representado por una matriz.*
(ii) *El operador unidad está representado por la matriz unidad.*

(iii) *Los operadores lineales reales están representados por matrices hermíticas.*

(iv) *ξ y las funciones de ξ están representados por matrices diagonales.*

(v) *La matriz que representa el producto de dos operadores lineales es igual al producto de las matrices que representan los dos factores.*

Consideremos ahora el caso de un solo observable ξ con autovalores continuos. El representante de α, $\langle\xi'|\alpha|\xi''\rangle$ es en este caso una función de dos variables ξ' y ξ'' que pueden variar continuamente. Es muy conveniente llamar a esta función 'matriz', generalizando el sentido de esta palabra, para así poder utilizar la misma terminología en el caso discreto y en el continuo. Por supuesto, tales matrices generalizadas ya no pueden escribirse, como ocurre con las matrices ordinarias, en una disposición bidimensional, ya que el número de filas y columnas es una infinidad del mismo orden que el número de puntos de una recta y el número de sus elementos una infinidad del mismo orden que el número de puntos de una superficie.

Daremos nuestras definiciones para estas matrices generalizadas de forma que sigan valiendo para el caso continuo las reglas (i)-(v) del caso discreto. El operador unidad está representado por $\delta(\xi'-\xi'')$ y la matriz generalizada formada por estos elementos será por definición la *matriz unidad*. La ecuación (41) seguirá siendo la condición para que α sea real, y por definición, diremos que una matriz generalizada es hermítica cuando sus elementos $\langle\xi'|\alpha|\xi''\rangle$ cumplan dicha condición. ξ está representado por

$$\langle\xi'|\xi|\xi''\rangle = \xi'\delta(\xi'-\xi'') \tag{45}$$

y $f(\xi)$ por

$$\langle\xi'|f(\xi)|\xi''\rangle = f(\xi')\delta(\xi'-\xi''), \tag{46}$$

y las matrices generalizadas formadas por tales elementos diremos que son *matrices diagonales*. Según (11) hubiéramos podido escribir también ξ'' y $f(\xi'')$ como coeficientes de las $\delta(\xi'-\xi'')$ en los segundos miembros de (45) y (46) respectivamente. Usando (24) obtenemos la ecuación correspondiente a (44)

$$\langle\xi'|\alpha\beta|\xi''\rangle = \int \langle\xi'|\alpha|\xi'''\rangle\, d\xi'''\, \langle\xi'''|\beta|\xi''\rangle, \tag{47}$$

en que figura una integral en vez de una suma, y diremos que la matriz formada por los elementos del segundo miembro es por definición el producto de las matrices formadas por $\langle\xi'|\alpha|\xi''\rangle$ y $\langle\xi'|\beta|\xi''\rangle$. Con estas definiciones logramos un completo paralelismo entre el caso discreto y el continuo, de forma que las reglas (i)-(v) valen para ambos casos.

Surge el problema de definir cuándo una matriz general es diagonal, puesto que hasta ahora sólo hemos dicho que los segundos miembros de (45) y (46) son ejemplos de matrices diagonales. Podríamos de suyo inclinarnos

a definir como diagonal cualquier matriz cuyos elementos (ξ', ξ'') fueran todos nulos salvo los correspondientes a valores de ξ' que difieren infinitamente poco de ξ''. Pero esta definición no sería satisfactoria, pues en el caso discreto las matrices diagonales gozan de una propiedad importante, la de conmutar siempre entre sí, y queremos que dicha propiedad siga valiendo también para el caso continuo. Para que en el caso continuo la matriz formada por los elementos $\langle\xi'|\omega|\xi''\rangle$ conmute con la formada por los elementos del segundo miembro de (45), ha de verificarse, según la regla de multiplicación (47)

$$\int \langle\xi'|\omega|\xi'''\rangle\, d\xi'''\, \xi'''\, \delta(\xi'''-\xi'') = \int \xi'\, \delta(\xi'-\xi''')\, d\xi'''\, \langle\xi'''|\omega|\xi''\rangle.$$

Usando la fórmula (4), la anterior se reduce a

$$\langle\xi'|\omega|\xi''\rangle\xi'' = \xi'\langle\xi'|\omega|\xi''\rangle \tag{48}$$

o sea

$$(\xi'-\xi'')\langle\xi'|\omega|\xi''\rangle = 0.$$

Y según la regla por la que de (12) dedujimos (13) esto equivale a

$$\langle\xi'|\omega|\xi''\rangle = c'\,\delta(\xi'-\xi'')$$

donde c' es un número que puede depender de ξ'. Luego, $\langle\xi'|\omega|\xi''\rangle$ es de la forma del segundo miembro de (46). Por esta razón diremos que *únicamente las matrices cuyos elementos son de la forma del segundo miembro de* (46) *son por definición matrices diagonales.* Es fácil comprobar que tales matrices conmutan todas entre sí. Se pueden formar otras matrices cuyos elementos (ξ', ξ'') sean todos nulos cuando ξ' difiera apreciablemente de ξ'' y que tengan otra forma de singularidad para ξ' igual a ξ'' [más adelante introduciremos la derivada $\delta'(x)$ de la función δ, y $\delta'(\xi'-\xi'')$ constituirá un ejemplo de ello; véase § ecuación (19)], pero tales matrices no serán diagonales según nuestra definición.

Pasemos ahora al caso de un solo ξ con autovalores de ambas clases, discretos y continuos. Llamemos ξ^r, ξ^s a los autovalores discretos y ξ', ξ'' a los continuos. El representante de α constará ahora de cuatro clases de cantidades: $\langle\xi^r|\alpha|\xi^s\rangle$, $\langle\xi^r|\alpha|\xi'\rangle$, $\langle\xi'|\alpha|\xi^r\rangle$, $\langle\xi'|\alpha|\xi''\rangle$. Podemos considerar que el conjunto de estas cantidades forman un tipo de matriz más general con ciertas filas y columnas discretas e intervalos continuos de filas y columnas. Para este tipo de matrices más general definiremos también la matriz unidad, los conceptos de matriz hermítica y de matriz diagonal y el producto de matrices, de forma que sigan siendo válidas las reglas (i)-(v). Los detalles son una generalización inmediata de lo que acabamos de hacer, y no vale la pena explicitarlos.

Volvamos de nuevo al caso general de varios observables $\xi_1, \xi_2, \ldots \xi_u$. El representante de α, dado por la expresión (39), puede considerarse también como una matriz, cuyas filas corresponden a diferentes valores de

$\xi'_1, \ldots, \xi'_u$ y las columnas a diferentes valores de $\xi''_1 \ldots, \xi''_u$. Salvo en el caso de que todos los ξ tengan únicamente autovalores discretos, esta matriz será del tipo generalizado con intervalos continuos de filas y columnas. Podemos modificar nuestras definiciones de forma que las reglas (i)-(v) sigan siendo válidas, si generalizamos la regla (iv) del modo siguiente:

(iv) *Cada uno de los* ξ_m $(m = 1, 2, \ldots, u)$ *y todas las funciones de ellos están representados por matrices diagonales.*

En este caso, diremos que una matriz es diagonal cuando su elemento general $\langle \xi'_1 \ldots \xi'_u | \omega | \xi''_1 \ldots \xi''_u \rangle$ sea de la forma

$$\langle \xi'_1 \ldots \xi'_u | \omega | \xi''_1 \ldots \xi''_u \rangle = c' \, \delta_{\xi'_1 \xi''_1} \ldots \delta_{\xi'_v \xi''_v} \delta(\xi'_{v+1} - \xi''_{v+1}) \ldots \delta(\xi'_u - \xi''_u) \qquad (49)$$

donde suponemos que $\xi_1, \ldots, \xi_v$ tienen autovalores discretos y $\xi_{v+1}, \ldots, \xi_u$ los tienen continuos, y siendo c' una función cualquiera de los ξ'. Esta definición es una generalización de la dada para el caso de un solo ξ, y de ella se deduce que las matrices diagonales siempre conmutan entre sí. Las demás definiciones son inmediatas y no vale la pena explicitarlas.

De este modo, todo operador lineal está representado por una matriz. La suma de dos operadores lineales está representada por la matriz suma de las que representan ambos operadores, lo que junto con la regla (v) nos dice que *las matrices están sujetas a las mismas operaciones algebraicas que los operadores.* Cualquier ecuación algebraica válida entre operadores lineales también será válida entre las matrices que representan a dichos operadores.

Podemos ampliar nuestro esquema de matrices para que abarque también a los representantes de los vectores bra y ket. Todas las matrices que representan operadores lineales son matrices cuadradas, con el mismo número de filas que de columnas, o mejor dicho, existe una correspondencia biyectiva entre sus filas y sus columnas. Podemos considerar el representante de un ket $|P\rangle$ como *una matriz de una sola columna,* obtenida colocando uno debajo de otro todos los números $\langle \xi'_1 \ldots \xi'_u | P \rangle$ que forman dicho representante. El número de filas de esta matriz será igual al número de filas o de columnas de las matrices cuadradas que representan operadores lineales. Tal matriz de una columna puede multiplicarse a su izquierda por la matriz cuadrada $\langle \xi'_1 \ldots \xi'_u | \alpha | \xi''_1 \ldots \xi''_u \rangle$ que representa un operador lineal, mediante una regla análoga a la de multiplicación de matrices cuadradas. El producto es una matriz también de una sola columna cuyos elementos vienen dados por

$$\sum_{\xi''_1 \ldots \xi''_v} \int \ldots \int \langle \xi'_1 \ldots \xi'_u | \alpha | \xi''_1 \ldots \xi''_u \rangle \, d\xi''_{v+1} \ldots d\xi''_u \, \langle \xi''_1 \ldots \xi''_u | P \rangle.$$

Según (35) esto es precisamente igual a $\langle \xi'_1 \ldots \xi'_u | \alpha | P \rangle$, o sea al representante de $\alpha | P \rangle$. Análogamente podemos considerar el representante de un

bra $\langle Q|$ como *una matriz de una sola fila*, obtenida colocando uno al lado de otro todos los números $\langle Q|\xi'_1 \ldots \xi'_u\rangle$. Tal matriz de una fila puede multiplicarse a la derecha por una matriz cuadrada $\langle \xi'_1 \ldots \xi'_u|\alpha|\xi''_1 \ldots \xi''_u\rangle$, y el producto será otra matriz de una sola fila que es precisamente igual al representante de $\langle Q|\alpha$. La matriz de una sola fila que representa a $\langle Q|$ puede multiplicarse a la derecha por la matriz de una sola columna que representa a $|P\rangle$, y el producto será una matriz de un solo elemento que vale $\langle Q|P\rangle$. Finalmente, la matriz de una sola fila que representa a $\langle Q|$ puede multiplicarse a la izquierda por la matriz de una sola columna que representa a $|P\rangle$, y el producto será una matriz cuadrada que es precisamente igual al representante de $|P\rangle\langle Q|$. De este modo todos nuestros símbolos abstractos — operadores lineales, vectores bra y vectores ket — pueden representarse por matrices que están sujetas a las mismas operaciones algebraicas que los correspondientes símbolos abstractos.

18. *Amplitudes de probabilidad*

Las representaciones son muy importantes en la interpretación física de la mecánica cuántica, ya que nos proporcionan un método práctico de obtener las probabilidades de que los observables tengan valores dados. En § 12 obtuvimos la probabilidad de que un observable tenga un valor determinado cualquiera para un estado dado, y en § 13, generalizando este resultado, obtuvimos la probabilidad de que un conjunto de observables que conmutan tengan simultáneamente unos valores determinados para un estado dado. Apliquemos ahora este resultado a un conjunto completo de observables que conmutan, es decir, al conjunto de los ξ que venimos estudiando. Según la fórmula (51) de § 13, la probabilidad de que para el estado correspondiente al ket normalizado $|x\rangle$ cada uno de los ξ_r tenga el valor ξ'_r es

$$P_{\xi'_1 \ldots \xi'_u} = \langle x|\delta_{\xi_1\xi'_1}\delta_{\xi_2\xi'_2}\ldots\delta_{\xi_u\xi'_u}|x\rangle. \tag{50}$$

Si todos los ξ tienen autovalores discretos, podemos emplear (35) con $v = u$ y sin integrales, con lo que resulta

$$\begin{aligned} P_{\xi'_1 \ldots \xi'_u} &= \sum_{\xi''_1 \ldots \xi''_u} \langle x|\delta_{\xi_1\xi'_1}\ \delta_{\xi_2\xi'_2}\ldots\ \delta_{\xi_u\xi'_u}|\xi''_1 \ldots \xi''_u\rangle\langle\xi''_1 \ldots \xi''_u|x\rangle \\ &= \sum_{\xi''_1 \ldots \xi''_u} \langle x|\delta_{\xi''_1\xi'_1}\ \delta_{\xi''_2\xi'_2}\ldots\ \delta_{\xi''_u\xi'_u}|\xi''_1 \ldots \xi''_u\rangle\langle\xi''_1 \ldots \xi''_u|x\rangle \\ &= \langle x|\xi'_1 \ldots \xi'_u\rangle\langle\xi'_1 \ldots \xi'_u|x\rangle \\ &= |\langle\xi'_1 \ldots \xi'_u|x\rangle|^2. \end{aligned} \tag{51}$$

Así obtenemos el resultado sencillo que podemos enunciar del siguiente modo: *la probabilidad de que para un cierto estado los ξ tengan valores ξ' es igual al cuadrado del módulo de la coordenada correspondiente del ket normalizado relativo al estado en cuestión.*

Si no todos los ξ tienen autovalores discretos, y por ejemplo $\xi_1, \ldots \xi_v$ tienen autovalores discretos y $\xi_{v+1}, \ldots, \xi_u$ los tienen continuos, entonces para que el resultado tenga sentido físico, hemos de calcular la probabilidad de que cada ξ_r $(r = 1, \ldots, v)$ tenga un valor determinado ξ'_r y cada ξ_s $(s = v + 1, \ldots, u)$ tenga valores comprendidos en un pequeño intervalo determinado $(\xi'_s, \xi'_s + d\xi'_s)$. Para ello hemos de reemplazar cada factor $\delta_{t_s t'_s}$ que figura en (50) por un factor χ_s, que es la función del observable ξ_s que vale uno cuando ξ_s toma valores comprendidos en el intervalo $(\xi'_s\ \xi'_s + d\xi'_s)$, y cero para cualquier otro caso. Operando como antes a partir de (35) obtendremos para dicha probabilidad el valor

$$P_{t'_1 \ldots t'_u} d\xi'_{v+1} \ldots d\xi'_u = |\langle \xi'_1 \ldots \xi'_u | x \rangle|^2 \, d\xi'_{v+1} \ldots d\xi'_u . \tag{52}$$

Así pues, en todos los casos *la distribución de probabilidad de los valores de los ξ para un cierto estado viene dada por el cuadrado del módulo del representante del ket normalizado correspondiente a dicho estado.*

Con ello queda justificado que a los números que constituyen el representante de un ket (o bra) normalizado se les denomine *amplitudes de probabilidad.* El cuadrado del módulo de una amplitud de probabilidad es una probabilidad ordinaria, o una probabilidad por unidad de intervalo para aquellas variables que toman valores en intervalos continuos.

Puede ocurrir que nos interese un estado cuyo ket correspondiente $|x\rangle$ no pueda normalizarse. Por ejemplo, cuando el estado es un autoestado de un observable que pertenece a un autovalor comprendido en un intervalo continuo de autovalores. Las fórmulas (51) y (52) siguen siendo válidas en este caso, y darán la probabilidad relativa de que los ξ tengan valores determinados o tengan valores comprendidos en pequeños intervalos prefijados, es decir, nos darán correctamente la relación entre las probabilidades para los distintos ξ'. Los números $\langle \xi'_1 \ldots \xi'_u | x \rangle$ se denominan en este caso *amplitudes de probabilidad relativa.*

La representación en la que son válidos los resultados anteriores se caracteriza por tener unos vectores básicos que son autovalores comunes a todos los ξ. Puede caracterizarse también por la existencia de que cada uno de los ξ esté representado por una matriz diagonal, ya que puede verse fácilmente que esta condición equivale a la anterior. Este segundo modo de caracterizarla es de ordinario el más práctico, y lo formularemos brevemente diciendo que cada uno de los ξ '*es diagonal en la representación*'.

Si los ξ forman un conjunto completo de observables que conmutan, la representación queda completamente determinada por esta caracterización, a menos de un factor de fase arbitrario en los vectores básicos. Cada bra

básico $\langle\xi'_1 \ldots \xi'_u|$ puede multiplicarse por $e^{i\gamma'}$ — siendo γ' cualquier función real de las variables $\xi'_1, \ldots, \xi'_u$ — sin que cambie por ello ninguna de las condiciones que ha de satisfacer la representación, es decir, la condición de que los ξ sean diagonales o la de que los vectores básicos sean autovectores comunes a los ξ y las propiedades fundamentales de los vectores básicos (34) y (35). Cuando cambiamos los vectores básicos de este modo, el representante $\langle\xi'_1 \ldots \xi'_u|P\rangle$ de un ket $|P\rangle$ queda multiplicado por $e^{i\gamma'}$, el representante $\langle Q|\xi'_1 \ldots \xi'_u|$ de un bra $\langle Q|$ queda multiplicado por $e^{-i\gamma'}$ y el representante $\langle\xi'_1 \ldots \xi'_u|\alpha|\xi''_1 \ldots \xi''_u\rangle$ de un operador lineal α queda multiplicado por $e^{(i\gamma'-\gamma'')}$. Evidentemente, las probabilidades o probabilidades relativas (51) y (52) quedan inalteradas.

Las probabilidades que se calculan en los problemas prácticos de mecánica cuántica se obtienen casi siempre a partir de los cuadrados de los módulos de amplitudes de probabilidad o amplitudes de probabilidad relativa. Incluso en el caso de que sólo interese la probabilidad de que un conjunto incompleto de observables que conmutan tengan valores determinados, de ordinario es necesario construir primero un conjunto completo introduciendo para ello nuevos observables que conmutan. A continuación obtener la probabilidad de que el conjunto completo tenga determinados valores (elevando al cuadrado el módulo de la amplitud de probabilidad), y después sumar o integrar estas probabilidades para todos los valores posibles de los observables suplementarios. De ordinario no es viable una aplicación más directa de la fórmula (51) de § 13.

En la práctica, para introducir una representación

(i) buscamos observables que convendría que fueran diagonales, bien porque nos interesa conocer sus probabilidades o bien por razones de simplificación matemática;
(ii) hemos de comprobar si todos ellos conmutan (condición necesaria, ya que las matrices diagonales siempre conmutan);
(iii) comprobamos si forman un conjunto completo, y si no es así, añadimos más observables que conmutan hasta formar un conjunto completo de observables que conmutan;
(iv) establecemos una representación ortogonal en la que este conjunto completo de observables sea diagonal.

La representación estará entonces completamente determinada a menos de los factores de fase arbitrarios. En la mayor parte de los casos estos factores de fase arbitrarios carecen de importancia y son triviales, de forma que podemos considerar la representación como completamente determinada por los observables que son diagonales en ella. Este hecho lo hemos supuesto ya en nuestra notación al escribir en los representantes unas letras que indican qué observables son diagonales en ella, como única indicación de la representación a la que pertenecen.

Puede ocurrir que nos interese utilizar dos representaciones para un mismo sistema dinámico. Supongamos que el conjunto completo de observables que conmutan y que son diagonales sea en una de ellas $\xi_1, \ldots, \xi_u$, y en la otra $\eta_1, \ldots, \eta_w$. Los bras básicos de la primera representación serán $\langle \xi'_1 \ldots \xi'_u |$, y los de la segunda $\langle \eta'_1 \ldots \eta'_w |$. Un ket $|P\rangle$ tendrá ahora dos representantes: $\langle \xi'_1 \ldots \xi'_u | P\rangle$ y $\langle \eta'_1 \ldots \eta'_w | P\rangle$. Si $\xi_1, \ldots, \xi_v$ tienen autovalores discretos y $\xi_{v+1}, \ldots, \xi_u$ los tienen continuos, y si $\eta_1, \ldots, \eta_x$ tienen autovalores discretos y $\eta_{x+1}, \ldots, \eta_w$ los tienen continuos, de (35) obtenemos

$$\langle \eta'_1 \ldots \eta'_w | P\rangle = \sum_{\xi'_1 \ldots \xi'_v} \int \cdots \int \langle \eta'_1 \ldots \eta'_w | \xi'_1 \ldots \xi'_u \rangle \, d\xi'_{v+1} \ldots d\xi'_u \, \langle \xi'_1 \ldots \xi'_u | P\rangle, \quad (53)$$

e intercambiando las ξ y las η

$$\langle \xi'_1 \ldots \xi'_u | P\rangle = \sum_{\eta'_1 \ldots \eta'_x} \int \cdots \int \langle \xi'_1 \ldots \xi'_u | \eta'_1 \ldots \eta'_w \rangle \, d\eta'_{x+1} \ldots d\eta'_w \, \langle \eta'_1 \ldots \eta'_w | P\rangle. \quad (54)$$

Estas son las ecuaciones de transformación que nos dan un representante de $|P\rangle$ en función del otro. En ella vemos que cada representante se puede expresar linealmente en función del otro, con los coeficientes

$$\langle \eta'_1 \ldots \eta'_w | \xi'_1 \ldots \xi'_u \rangle, \qquad \langle \xi'_1 \ldots \xi'_u | \eta'_1 \ldots \eta'_w \rangle \quad (55)$$

Estas cantidades se denominan *funciones de transformación*. Se pueden escribir ecuaciones análogas para relacionar los dos representantes de un vector bra o de un operador lineal. Las funciones de transformación (55) son el artificio que permite pasar en todos los casos de una representación a otra. Las funciones de transformación son complejas conjugadas unas de otras, y satisfacen las condiciones

$$\sum_{\xi'_1 \ldots \xi'_v} \int \cdots \int \langle \eta'_1 \ldots \eta'_w | \xi'_1 \ldots \xi'_u \rangle \, d\xi'_{v+1} \ldots d\xi'_u \, \langle \xi'_1 \ldots \xi'_u | \eta''_1 \ldots \eta''_w \rangle$$
$$= \delta_{\eta'_1 \eta''_1} \ldots \delta_{\eta'_x \eta''_x} \delta(\eta'_{x+1} - \eta''_{x+1}) \ldots \delta(\eta'_w - \eta''_w) \quad (56)$$

y las condiciones correspondientes obtenidas por intercambio de las ξ con los η, como puede comprobarse fácilmente a partir de (35) y (34) y las correspondientes ecuaciones para los η.

Las funciones de transformación son ejemplos de amplitudes de probabilidad o amplitudes de probabilidad relativa. Consideremos el caso en que tanto los ξ como los η tienen únicamente autovalores discretos. En este caso, el ket básico $|\eta'_1 \ldots \eta'_w\rangle$ está normalizado, y su representante en la representación ξ, $\langle \xi'_1 \ldots \xi'_u | \eta'_1 \ldots \eta'_w \rangle$, es una amplitud de probabilidad para cada conjunto de valores de las ξ'. El estado al que se refieren dichas amplitudes de probabilidad, o sea el estado correspondiente a $|\eta'_1 \ldots \eta'_w\rangle$, está caracterizado por la condición de que una medida simultánea de los

observables $\eta_1, \ldots, \eta_w$ dé con seguridad los resultados $\eta'_1, \ldots, \eta'_w$. En consecuencia, $|\langle\xi'_1\ldots\xi'_u|\eta'_1\ldots\eta'_w\rangle|^2$ es la probabilidad de que los ξ tengan los valores $\xi'_1, \ldots, \xi'_u$ para el estado en que los η tienen con seguridad los valores $\eta'_1, \ldots, \eta'_w$. Y puesto que

$$|\langle\xi'_1\ldots\xi'_u|\eta'_1\ldots\eta'_w\rangle|^2 = |\langle\eta'_1\ldots\eta'_w|\xi'_1\ldots\xi'_u\rangle|^2,$$

tenemos el siguiente teorema de reciprocidad: *la probabilidad de que los ξ tengan los valores ξ' para un estado en el que los η tienen ciertamente los valores η', es igual a la probabilidad de que los η tengan los valores η' para un estado en el que los ξ tienen ciertamente los valores ξ'.*

Si todos los η tienen autovalores discretos y algunos ξ tienen autovalores continuos, $|\langle\xi'_1\ldots\xi'_u|\eta'_1\ldots\eta'_w\rangle|^2$ sigue dándonos la distribución de probabilidad de los valores de los ξ para el estado en el que los η tienen ciertamente los valores η'. Si algunos de los η tienen autovalores continuos, $|\eta'_1\ldots\eta'_w\rangle$ no está normalizado y en este caso $|\langle\xi'_1\ldots\xi'_u|\eta'_1\ldots\eta'_w\rangle|^2$ únicamente nos da la distribución de probabilidad relativa de los valores de los ξ para el estado en el que los η tienen ciertamente los valores η'.

19. *Teoremas sobre funciones de observables*

Vamos a poner de manifiesto el valor matemático de las representaciones, utilizándolas para demostrar algunos teoremas.

TEOREMA 1. *Todo operador lineal que conmuta con un observable ξ conmuta también con cualquier función de ξ.*

El teorema es evidente si la función puede expresarse en serie de potencias. Para demostrarlo en el caso general, sea ω el operador lineal en cuestión, que verificará la ecuación

$$\xi\omega - \omega\xi = 0. \tag{57}$$

Introduzcamos una representación en la que ξ sea diagonal. Si ξ por sí solo no constituye un conjunto completo de observables, tendremos que añadir nuevos observables β a fin de tener un conjunto completo, y elegir entonces una representación en la que ξ y los β sean diagonales. (El caso en que ξ constituye por sí solo un conjunto completo de observables podemos considerarlo como un caso particular del anterior, donde el número de variables β es cero.) En esta representación, la ecuación (57) se convierte en

$$\langle\xi'\beta'|\xi\omega - \omega\xi|\xi''\beta''\rangle = 0,$$

que se reduce a

$$\xi'\langle\xi'\beta'|\omega|\xi''\beta''\rangle - \langle\xi'\beta'|\omega|\xi''\beta''\rangle\xi'' = 0.$$

Si los autovalores de ξ son discretos, esta ecuación muestra que todos los elementos $\langle\xi'\beta'|\omega|\xi''\beta''\rangle$ de la matriz de ω son nulos excepto aquellos para

los que $\xi' = \xi''$. Si los autovalores de ξ son continuos, la ecuación muestra —igual que (48)—que $\langle\xi'\beta'|\omega|\xi''\beta''\rangle$ es de la forma

$$\langle\xi'\beta'|\omega|\xi''\beta''\rangle = c\,\delta(\xi' - \xi'),$$

siendo c una cierta función de ξ', de las β' y de las β''. En ambos casos podemos decir que la matriz que representa ω *es diagonal respecto a* ξ'. Si llamamos $f(\xi)$ a una función cualquiera de ξ, de acuerdo con la teoría general de § 11, $f(\xi''')$ ha de estar definida para todo autovalor ξ''' de ξ, y tanto en uno como en otro caso podemos deducir

$$f(\xi')\langle\xi'\beta'|\omega|\xi''\beta''\rangle - \langle\xi'\beta'|\omega|\xi''\beta''\rangle f(\xi'') = 0.$$

De donde resulta

$$\langle\xi'\beta'|f(\xi)\omega - \omega f(\xi)|\xi''\beta''\rangle = 0,$$

o sea

$$f(\xi)\omega - \omega f(\xi) = 0$$

quedando así demostrado el teorema.

Como caso particular del teorema, llegamos a la conclusión de que todo observable que conmuta con un observable ξ conmuta también con cualquier función de ξ. Esta conclusión aparece como una necesidad física, si identificamos—como lo hicimos en § 13—la condición de conmutabilidad de dos observables con la condición de compatibilidad de las observaciones correspondientes. Cualquier observación compatible con la medida de un observable ξ ha de ser también compatible con la medida de $f(\xi)$, ya que toda medida de ξ incluye en sí misma la medida de $f(\xi)$.

Teorema 2. *Todo operador lineal que conmuta con cada uno de los observables de un conjunto completo es función de dichos observables.*

Sea ω el operador lineal en cuestión, y $\xi_1, \xi_2, \ldots, \xi_u$ el conjunto completo de observables que conmutan. Introduzcamos una representación en la que estos observables sean diagonales. Puesto que ω conmuta con cada uno de los ξ, su matriz representante es diagonal con respecto a cada uno de los ξ según el razonamiento anterior. Por tanto, dicha matriz es diagonal, y es de la forma (49) con un número c' función de las ξ'. Esta matriz es el representante de la función de los ξ que tenga la misma forma que la función c' de las ξ', de donde ω es igual a dicha función de los ξ.

Teorema 3. *Si un observable ξ y un operador lineal g son tales que todo operador lineal que conmuta con ξ conmuta también con g, entonces g es función de ξ.*

Este es el recíproco del teorema 1. Para demostrarlo emplearemos la misma representación del teorema 1, en la que ξ es diagonal. Observemos previamente que g conmuta con ξ y, por tanto, el representante de ω ha de ser diagonal respecto a ξ, es decir, ha de ser de la forma

$$\langle\xi'\beta'|g|\xi''\beta''\rangle = a(\xi'\beta'\beta'')\delta_{\xi'\xi''} \qquad \text{o bien} \qquad a(\xi'\beta'\beta'')\delta(\xi' - \xi''),$$

según tenga ξ autovalores discretos o continuos. Sea ω un operador lineal

cualquiera que conmuta con ξ cuyo representante será, en consecuencia, de la forma

$$\langle \xi'\beta'|\omega|\xi''\beta''\rangle = b(\xi'\beta'\beta'')\delta_{\xi'\xi''} \qquad \text{o bien} \qquad b(\xi'\beta'\beta'')\delta(\xi'-\xi'').$$

Por hipótesis ω conmuta también con g, luego

$$\langle \xi'\beta'|g\omega - \omega g|\xi''\beta''\rangle = 0. \tag{58}$$

Pongamos por caso que las β tengan autovalores discretos. A partir de (58), aplicando la ley de multiplicación de matrices, se llega a

$$\sum_{\beta'''} \{a(\xi'\beta'\beta''')b(\xi'\beta'''\beta'') - b(\xi'\beta'\beta''')a(\xi'\beta'''\beta'')\} = 0, \tag{59}$$

ya que el primer miembro de (58) es igual al primer miembro de (59) multiplicado por $\delta_{\xi'\xi''}$ o $\delta(\xi'-\xi'')$. La ecuación (59) ha de ser válida para toda función $b(\xi'\beta'\beta'')$. De ello deducimos que

$$a(\xi'\beta'\beta'') = 0 \quad \text{para} \quad \beta' \neq \beta'',$$
$$a(\xi'\beta'\beta') = a(\xi'\beta''\beta'').$$

El primero de estos resultados expresa que la matriz representante de g es diagonal y el segundo que $a(\xi'\beta'\beta')$ únicamente es función de ξ'. De aquí podemos deducir que g es la misma función de ξ que $a(\xi'\beta'\beta')$ es de ξ', quedando así demostrado el teorema. La demostración es análoga si algunos de los β tienen autovalores continuos.

Los teoremas 1 y 3 siguen siendo válidos si reemplazamos el observable ξ por un conjunto cualquiera de observables que conmutan $\xi_1, \xi_2, \ldots \xi_r$, pues para la demostración basta introducir algunos cambios formales.

cada bra $\langle \xi'_1|P\rangle, \langle \xi'_2|P\rangle \ldots, \langle \xi'_n|P\rangle$

20. *Nuevas formas de notación*

La teoría de representaciones que hemos desarrollado nos proporciona un procedimiento general para designar a los kets y a los bras. En una representación en que el conjunto completo de observables $\xi_1, \ldots, \xi_n$ sean diagonales, un ket $|P\rangle$ cualquiera tendrá por representante $\langle \xi'_1 \ldots \xi'_n|P\rangle$, que para abreviar escribiremos $\langle \xi'|P\rangle$. Este representante es una función definida de las variables ξ', que llamaremos $\psi(\xi')$. Así pues, la función ψ determina completamente el ket $|P\rangle$, y puede ser utilizada para designar a dicho ket en lugar de emplear el índice contrario P. Simbólicamente, si

$$\langle \xi'|P\rangle = \psi(\xi') \quad i)$$

escribiremos (60)

$$|P\rangle = |\psi(\xi)\rangle \quad ii)$$

Hemos de escribir $|P\rangle$ igual a $|\psi(\xi)\rangle$ y no a $|\psi(\xi')\rangle$, ya que no depende de

introducido ii) e i) $\langle \xi'|\psi(\xi)\rangle = \psi(\xi')$

y comparando con i)

$|P\rangle = |\psi(\xi)\rangle$

un conjunto particular de autovalores de los ξ sino sólo de la forma de la función ψ.

Si $f(\xi)$ es una función cualquiera de los observables $\xi_1, \ldots, \xi_u$, el ket $f(\xi)|P\rangle$ tendrá como representante

$$\langle\xi'|f(\xi)|P\rangle = f(\xi')\psi(\xi').$$

Y, por tanto, según (60), escribiremos

$$f(\xi)|P\rangle = |f(\xi)\psi(\xi)\rangle.$$

Utilizando la segunda de las ecuaciones (60), esto equivale a

$$f(\xi)|\psi(\xi)\rangle = |f(\xi)\psi(\xi)\rangle. \tag{61}$$

Este resultado es general, y es válido para cualesquiera f y ψ funciones de ξ. De ahí que en la nueva notación, la línea vertical | no sea ya necesaria, y ambos miembros de (61) los podemos escribir simplemente $f(\xi)\psi(\xi)\rangle$. Así la regla de la nueva notación será la siguiente:

si

$$\langle\xi'|P\rangle = \psi(\xi') \qquad \psi(\xi')$$

escribiremos

$$|P\rangle = \psi(\xi\ )\rangle \qquad \langle\xi'|\ \rangle = 1 \tag{62}$$

Incluso podremos abreviar $\psi(\xi)$ por ψ y sobreentender las variables ξ siempre que ello no dé lugar a confusión.

El ket $\psi(\xi)\rangle$ puede ser considerado como el resultado de aplicar el operador $\psi(\xi)$ a un ket que se escribe simplemente $\rangle$, sin índice ninguno. Al ket $\rangle$ le llamaremos *ket standard*. Todo ket puede expresarse como el producto de una función de las ξ por el ket standard. Por ejemplo, si en (62) tomamos como $|P\rangle$ el ket básico $|\xi''\rangle$, obtendremos

$$|\xi''\rangle = \delta_{\xi_1\xi_1''}\ldots\delta_{\xi_v\xi_v''}\delta(\xi_{v+1} - \xi''_{v+1})\ldots\delta(\xi_u - \xi''_u)\rangle \tag{63}$$

donde suponemos que los $\xi_1, \ldots, \xi_v$ tienen autovalores discretos y los $\xi_{v+1}, \ldots, \xi_u$ los tienen continuos. El ket standard se caracteriza por la condición de que su representante $\langle\xi'|\rangle$ es igual a uno para todo el dominio de las variables ξ', según puede verse haciendo $\psi = 1$ en (62).

Aún podemos abreviar más nuestra notación suprimiendo el símbolo $\rangle$ y sobreentendiendo el ket standard. En este caso, el ket se designa simplemente $\psi(\xi)$, como una función de los observables ξ. Una función de las ξ usada de este modo para caracterizar un ket se denomina *función de onda.* * El sistema de notación que emplea funciones de onda es el que acostumbran

* Este nombre se debe a que en los primeros tiempos de la mecánica cuántica todos los ejemplos que aparecían de esas funciones eran de forma ondulatoria. Pero desde el punto de vista de la moderna teoría general, dicho nombre no resulta ya descriptivo.

$\psi(\xi) = \psi(\xi)\rangle$ = fc de onda

a utilizar la mayoría de los autores en los cálculos de mecánica cuántica. Cuando lo usemos habremos de recordar que junto a cada función de onda se sobreentiende siempre un ket standard que la multiplica a su derecha, que nos prohíbe multiplicarla por un operador a su derecha. *Las funciones de onda sólo se pueden multiplicar por operadores a la izquierda.* Esto las distingue de las funciones ordinarias, que son operadores y pueden multiplicarse por otros operadores tanto a la izquierda como a la derecha. Una función de onda es precisamente el representante de un ket expresado como función de los observables ξ en vez de los autovalores ξ' de dichos observables. El cuadrado de su módulo da la probabilidad (o la probabilidad relativa si no está normalizado) de que, para el estado correspondiente, los ξ tengan valores dados o estén comprendidos en pequeños intervalos dados.

Para los bras se puede desarrollar la nueva notación del mismo modo que para los kets. Un bra $\langle Q|$ cuyo representante $\langle Q|\xi'\rangle$ sea $\phi(\xi')$, lo escribiremos $\langle\phi(\xi)|$. Con esta notación, el imaginario conjugado de $|\psi(\xi)\rangle$ es $\langle\psi(\xi)|$. Por tanto, la regla que hemos empleado hasta ahora según la cual designábamos un ket y su bra imaginario conjugado los dos con el mismo índice, habremos de generalizarla de este modo: *si para designar un ket se emplean números o funciones complejas, para designar su bra imaginario conjugado se habrán de emplear los números o funciones complejos conjugados.* Como en el caso de los kets, podemos comprobar que $\langle\phi(\xi)|f(\xi)$ y $\langle\phi(\xi)f(\xi)$ representan el mismo bra y que, en consecuencia, la línea vertical puede omitirse. Podemos considerar $\langle\phi(\xi)$ como el producto del operador lineal $\phi(\xi)$ por el *bra standard* $\langle$, que es el imaginario conjugado del ket standard $\rangle$ Podremos dejar sobreentendido el bra standard, y así un bra general se escribirá $\phi(\xi)$, que es el complejo conjugado de una función de onda. El complejo conjugado de una función de onda puede ser multiplicado por un operador lineal a la derecha, pero nunca por la izquierda. podemos construir productos triples de la forma $\langle f(\xi)\rangle$. Un producto triple de esta forma es un número igual a la suma o la integral de $f(\xi)$ para todo el dominio de autovalores de los ξ, o sea

$$\langle f(\xi)\rangle = \sum_{\xi'_1 \ldots \xi'_v} \int \ldots \int f(\xi')\, d\xi'_{v+1} \ldots d\xi'_u \tag{64}$$

para el caso en que los $\xi_1, \ldots, \xi_v$ tengan autovalores discretos, y los $\xi_{v+1}, \ldots, \xi_u$ tengan autovalores continuos.

Tanto el ket como el bra standard se definen siempre respecto a una representación. Si repitiéramos las consideraciones anteriores con una representación diferente en la que el conjunto completo de observables que es diagonal fuera η, o si cambiamos simplemente los factores de fase de la representación en la que los ξ son diagonales, obtendríamos un bra y un ket standard diferentes de los anteriores. Cuando trabajemos a la vez con

varios kets o bras standards diferentes habrá que distinguirlos mediante índices convenientes.

Vamos a discutir ahora un nuevo desarrollo de nuestra notación de gran importancia para tratar sistemas dinámicos complicados. Sea un sistema dinámico que puede describirse en función de ciertas variables dinámicas susceptibles de clasificarse en dos conjuntos — conjunto A y conjunto B — tales que cualquier elemento de A conmuta con cualquier elemento de B. Una variable dinámica general ha de expresarse como una función que depende a la vez de las variables A y de las variables B. Consideremos otro sistema dinámico cuyas únicas variables dinámicas sean las variables A y llamémosle sistema A. Y análogamente consideremos un tercer sistema dinámico, el sistema B, cuyas únicas variables dinámicas sean las variables B. El sistema original puede considerarse como una combinación de los sistemas A y B según un esquema matemático que damos a continuación.

Sea $|a\rangle$ un ket cualquiera del sistema A y $|b\rangle$ un ket cualquiera del sistema B. Establezcamos un producto $|a\rangle|b\rangle$ que disfrute de las propiedades conmutativa y distributiva de la multiplicación:

$$
\begin{aligned}
|a\rangle|b\rangle &= |b\rangle|a\rangle, \\
\{c_1|a_1\rangle+c_2|a_2\rangle\}|b\rangle &= c_1|a_1\rangle|b\rangle+c_2|a_2\rangle|b\rangle, \\
|a\rangle\{c_1|b_1\rangle+c_2|b_2\rangle\} &= c_1|a\rangle|b_1\rangle+c_2|a\rangle|b_2\rangle,
\end{aligned}
$$

siendo las c números cualesquiera. Para cada variable A podemos definir su acción sobre el producto $|a\rangle|b\rangle$ estableciendo que actúa sólo sobre el factor $|a\rangle$ y conmuta con el factor $|b\rangle$. Análogamente podemos definir la acción de cada variable B sobre el producto estableciendo que opera sólo sobre el factor $|b\rangle$ y conmuta con el $|a\rangle$. (Lo que hace que toda variable A conmute con toda variable B.) De este modo cada variable del sistema original puede operar sobre el producto $|a\rangle|b\rangle$, de forma que este producto puede escribirse $|ab\rangle$, pues los dos índices a y b son suficientes para especificarlo. De este modo obtenemos las ecuaciones fundamentales

$$|a\rangle|b\rangle = |b\rangle|a\rangle = |ab\rangle. \tag{65}$$

Esta multiplicación es de un tipo diferente a todas las que hemos utilizado hasta aquí en la teoría. Los vectores ket $|a\rangle$ y $|b\rangle$ están en dos espacios vectoriales diferentes, y su producto está en un tercer espacio vectorial, que podemos llamar producto de los espacios vectoriales. El número de dimensiones del espacio producto es igual al producto de los números de dimensiones de los espacios factores. Un vector ket del espacio producto no es en general de la forma (65), sino una suma o integral de kets de esta forma.

Consideremos una representación del sistema A en la que sean diagonales un conjunto completo de observables que conmutan ξ_A. Tendremos así para el sistema A los bras básicos $\langle\xi'_A|$. Análogamente, en una represen-

tación del sistema B en la que los observables ξ_B son diagonales, tendremos para el sistema B los bras básicos $\langle\xi'_B|$. Los productos

$$\langle\xi'_A|\langle\xi'_B| = \langle\xi'_A\xi'_B| \tag{66}$$

nos proporcionarán los bras básicos de una representación del sistema original en la que todos los ξ_A y ξ_B serán diagonales. Los ξ_A y ξ_B juntos formarán un conjunto completo de observables que conmutan para el sistema original. A partir de (65) y (66) resulta

$$\langle\xi'_A|a\rangle\langle\xi'_B|b\rangle = \langle\xi'_A\xi'_B|ab\rangle, \tag{67}$$

lo que nos muestra que el representante de $|ab\rangle$ es igual al producto de los representantes de $|a\rangle$ y $|b\rangle$ en sus respectivas representaciones.

Podemos introducir el ket standard $\rangle_A$ correspondiente al sistema A y a la representación en que los ξ_A son diagonales, y asimismo el ket standard $\rangle_B$ correspondiente al sistema B y a la representación en que los ξ_B son diagonales. Su producto $\rangle_A\rangle_B$ es entonces el ket standard correspondiente al sistema original y a la representación en que los ξ_A y los ξ_B son diagonales. Un ket cualquiera del sistema original se puede expresar en la forma

$$\psi(\xi_A\xi_B)\rangle_A\rangle_B. \tag{68}$$

Pudiera ocurrir que en ciertos cálculos nos interese emplear una representación particular del sistema B — por ejemplo la representación anterior en la que los ξ_B son diagonales — pero que en cambio no nos interesa introducir ninguna representación particular del sistema A. Nos convendrá entonces usar el ket standard $\rangle_B$ del sistema B y no introducir ningún ket standard del sistema A. En este caso todo ket del sistema original podremos escribirlo en la forma

$$|\xi_B\rangle\rangle_B, \tag{69}$$

donde $|\xi_B\rangle$ es un ket del sistema A y es además función de los ξ_B; es decir, es un ket del sistema A para cada conjunto de valores de los ξ_B. De hecho (69) es igual a (68) si hacemos

$$|\xi_B\rangle = \psi(\xi_A\xi_B)\rangle_A.$$

En (69) podemos dejar sobreentendido el ket standard $\rangle_B$, y entonces el ket general del sistema original nos aparece en la forma $|\xi_B\rangle$, que es un ket del sistema A y una función de onda de las variables ξ_B del sistema B. En §§ 66 y 79 hay ejemplos de esta notación.

Las consideraciones anteriores las podemos generalizar inmediatamente a un sistema dinámico que pueda describirse en función de variables dinámica susceptibles de clasificarse en tres o más conjuntos A, B, C, ... tales

que cualquier elemento de un conjunto conmuta con cualquier elemento de otro. La generalización de la ecuación (65) será

$$|a\rangle|b\rangle|c\rangle\ldots = |abc\ldots\rangle,$$

donde los factores del primer miembro son kets de los sistemas componentes, y el ket del segundo miembro es un ket del sistema original. De un modo semejante se pueden generalizar para este caso de varios factores las ecuaciones (66), (67) y (68).

IV

CONDICIONES CUÁNTICAS

21. *Corchetes de Poisson*

Hasta aquí nuestra labor ha consistido en construir un esquema matemático general que relaciona los estados y los observables en mecánica cuántica. Uno de los rasgos más característicos de este esquema es que en él los observables y las variables dinámicas en general son cantidades que no disfrutan de la propiedad conmutativa de la multiplicación. Ahora será necesario establecer ecuaciones que nos den el valor de $\xi\eta - \eta\xi$ para observables o variables dinámicas ξ y η cualesquiera, que sustituyan a la propiedad conmutativa de la multiplicación. Hasta que no conozcamos dichas ecuaciones no tendremos un esquema mecánico completo que pueda sustituir a la mecánica clásica. Estas nuevas ecuaciones reciben el nombre de *condiciones cuánticas* o *relaciones de conmutación.*

El problema de hallar condiciones cuánticas no tiene un carácter tan general como los que hemos considerado hasta ahora. Se trata de un problema que se nos presenta en particular cada vez que estudiamos un sistema dinámico concreto. Sin embargo, existe un método muy general para obtener condiciones cuánticas que se puede aplicar a una gran cantidad de sistemas dinámicos. Es el método de *analogía clásica,* que constituye el objeto de este capítulo. Los sistemas dinámicos a los que no se pueda aplicar dicho método deberemos estudiarlos individualmente y en cada caso tendremos que hacer consideraciones especiales.

La validez de la analogía clásica en el desarrollo de la mecánica cuántica proviene del hecho de que la mecánica clásica da una descripción correcta de los sistemas dinámicos en ciertas condiciones, a saber, cuando las partículas y los cuerpos que constituyen el sistema son suficientemente grandes para que la alteración que acompaña a la observación sea despreciable. Por tanto, la mecánica clásica tiene que ser un caso límite de la cuántica y, en consecuencia, debemos esperar que a conceptos importantes en mecánica clásica correspondan conceptos importantes en mecánica cuántica, y que tras un análisis de la naturaleza general de la analogía entre la mecánica clásica y la cuántica ha de ser posible deducir leyes y teoremas

de esta última que sean simples generalizaciones de resultados bien conocidos en la teoría clásica. En particular, podremos obtener las condiciones cuánticas como simples generalizaciones de la ley clásica según la cual todas las variables dinámicas conmutan.

Sea un sistema dinámico constituido por un conjunto de partículas en interacción. Para estudiar dicho sistema, podríamos tomar como variables dinámicas independientes las coordenadas cartesianas de cada partícula y las correspondientes componentes cartesianas de las velocidades. Pero es preferible utilizar las componentes del momento en lugar de las de la velocidad. Llamemos q_r a las coordenadas (r toma todos los valores desde 1 hasta tres veces el número total de partículas) y p_r a las correspondientes componentes del momento. Se dice que las q_r y las p_r son *coordenadas canónicas conjugadas.*

El método de las ecuaciones del movimiento de Lagrange permite introducir coordenadas q_r y momentos p_r de un modo más general, que también se puede aplicar a sistemas que no estén compuestos por partículas (por ejemplo, un sistema que contenga cuerpos rígidos). A dichas q_r y p_r más generales también se les llama coordenadas canónicas conjugadas. Toda variable dinámica se puede expresar como función de un conjunto de coordenadas canónicas cualesquiera.

En la teoría dinámica general es importante el concepto de *corchete de Poisson.* Todo par de variables dinámicas u y v tienen un P.B. (Poisson Bracket, o en castellano corchete de Poisson), que designaremos $[u, v]$, definido por

$$[u, v] = \sum_r \left\{ \frac{\partial u}{\partial q_r} \frac{\partial v}{\partial p_r} - \frac{\partial u}{\partial p_r} \frac{\partial v}{\partial q_r} \right\} \tag{1}$$

En la fórmula, a efectos de derivación, se consideran u y v expresadas en función de un conjunto de coordenadas canónicas q_r y p_r. El segundo miembro de (1) es independiente del conjunto de coordenadas canónicas que sirve de referencia, como se desprende de la definición general de coordenadas canónicas; luego el P.B. $[u, v]$ está bien definido.

Las principales propiedades de los corchetes de Poisson, que se deducen inmediatamente a partir de la definición (1) son

$$[u, v] = -[v, u], \tag{2}$$

$$[u, c] = 0, \tag{3}$$

en donde c es un número (que se puede considerar como un caso particular de variable dinámica).

$$\left.\begin{aligned} [u_1+u_2, v] &= [u_1, v]+[u_2, v], \\ [u, v_1+v_2] &= [u, v_1]+[u, v_2], \end{aligned}\right\} \tag{4}$$

$$[u_1 u_2, v] = \sum_r \left\{ \left(\frac{\partial u_1}{\partial q_r} u_2 + u_1 \frac{\partial u_2}{\partial q_r} \right) \frac{\partial v}{\partial p_r} - \left(\frac{\partial u_1}{\partial p_r} u_2 + u_1 \frac{\partial u_2}{\partial p_r} \right) \frac{\partial v}{\partial q_r} \right\}$$

$$\left.\begin{aligned} &= [u_1, v] u_2 + u_1 [u_2, v], \\ [u, v_1 v_2] &= [u, v_1] v_2 + v_1 [u, v_2]. \end{aligned}\right\} \quad (5)$$

Asimismo se puede comprobar fácilmente la identidad

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0 \quad (6)$$

Las ecuaciones (4) indican que el P.B. $[u, v]$ contiene a u y v linealmente, mientras que la (5) corresponde a las leyes ordinarias de derivación de un producto.

Intentemos introducir un P.B. en mecánica cuántica que tenga propiedades análogas al de la teoría clásica. Supondremos que el P.B. cuántico satisface todas las condiciones (1) a (6), pero teniendo en cuenta que ahora es necesario conservar el orden de los factores u_1 y u_2 en la primera de las ecuaciones (5) tal como lo hemos escrito aquí, y lo mismo con las v_1 y v_2 de la segunda ecuación (5). Estas condiciones son suficientes para determinar unívocamente la forma del P.B. en mecánica cuántica, como resulta del siguiente argumento. Podemos calcular el P.B. $[u_1 u_2, v_1 v_2]$ mediante dos procedimientos, según cual de las dos fórmulas (5) empleemos primero. Luego,

$$\begin{aligned} [u_1 u_2, v_1 v_2] &= [u_1, v_1 v_2] u_2 + u_1 [u_2, v_1 v_2] \\ &= \{[u_1, v_1] v_2 + v_1 [u_1, v_2]\} u_2 + u_1 \{[u_2, v_1] v_2 + v_1 [u_2, v_2]\} \\ &= [u_1, v_1] v_2 u_2 + v_1 [u_1, v_2] u_2 + u_1 [u_2, v_1] v_2 + u_1 v_1 [u_2, v_2] \end{aligned}$$

y

$$\begin{aligned} [u_1 u_2, v_1 v_2] &= [u_1 u_2, v_1] v_2 + v_1 [u_1 u_2, v_2] \\ &= [u_1, v_1] u_2 v_2 + u_1 [u_2, v_1] v_2 + v_1 [u_1, v_2] u_2 + v_1 u_1 [u_2, v_2]. \end{aligned}$$

Igualando ambos resultados obtenemos

$$[u_1, v_1](u_2 v_2 - v_2 u_2) = (u_1 v_1 - v_1 u_1)[u_2, v_2].$$

Como estos resultados tienen que verificarse para variables u_1 y v_1 completamente independientes de las u_2 y v_2, debe verificarse

$$\begin{aligned} u_1 v_1 - v_1 u_1 &= i\hbar [u_1, v_1], \\ u_2 v_2 - v_2 u_2 &= i\hbar [u_2, v_2], \end{aligned}$$

en donde $\hbar$ no puede depender ni de u_1 y v_1, ni de u_2 y v_2, y tiene que conmutar con $(u_1 v_1 - v_1 u_1)$. Luego $\hbar$ tiene que ser un número. Si queremos que, como ocurre en la teoría clásica, el P.B. de dos variables dinámicas reales sea real, según el final de la pág. 40, es necesario que $\hbar$ introducido así, con el coeficiente i, sea real. Llegamos así a la siguiente *definición cuántica del P.B.* $[u, v]$ *de dos variables* u *y* v *cualesquiera:*

$$uv - vu = i\hbar [u, v], \quad (7)$$

en la que $\hbar$ es una nueva constante universal que tiene las dimensiones de acción. Para que la teoría esté de acuerdo con la realidad experimental tenemos que elegir $\hbar$ igual a $h/2\pi$, siendo h la constante universal introducida por Planck, conocida con el nombre de constante de Planck. Puede comprobarse fácilmente que el P.B. de la mecánica cuántica verifica todas las condiciones (2), (3), (4), (5) y (6).

Con esto, el problema de hallar condiciones cuánticas se reduce a determinar los valores de los P.B. de la mecánica cuántica. La gran analogía que hay entre el P.B. de la mecánica cuántica definido por (7) y el P.B. de la mecánica clásica definido por (1) nos lleva a formular la hipótesis de que los P.B. de la primera, o por lo menos los más sencillos, tienen el mismo valor que los P.B. correspondientes de la mecánica clásica. Los P.B. más sencillos son los relativos a las propias variables canónicas, que en la teoría clásica tienen los siguientes valores:

$$\left.\begin{array}{c} [q_r, q_s] = 0, \qquad [p_r, p_s] = 0, \\ [q_r, p_s] = \delta_{rs}. \end{array}\right\} \quad (8)$$

Así supondremos que los P.B. correspondientes de la mecánica cuántica tienen también los valores dados por (8). Eliminando dichos P.B. con ayuda de (7), obtendremos las ecuaciones.

$$\left.\begin{array}{c} q_r q_s - q_s q_r = 0, \qquad p_r p_s - p_s p_r = 0, \\ q_r p_s - p_s q_r = i\hbar\delta_{rs}, \end{array}\right\} \quad (9)$$

que son las *condiciones cuánticas fundamentales*, que como vemos nos dan las relaciones de conmutación entre coordenadas canónicas conjugadas. A su vez, estas ecuaciones constituyen la base para calcular relaciones de conmutación entre otras variables dinámicas. Por ejemplo, si ξ y η son dos funciones de las q_r y de las p_r que se pueden desarrollar en serie de potencias, podemos expresar $\xi\eta - \eta\xi$ o $[\xi, \eta]$ en función de los P.B. fundamentales (8), aplicando reiteradamente las leyes (2), (3), (4) y (5), y calcular así dichas expresiones. En los casos sencillos el resultado es con frecuencia el mismo que el que se obtiene en la teoría clásica o bien difiere de él en que exige que los diversos factores de un producto estén en un orden dado, cosa que evidentemente desde el punto de vista clásico no tiene importancia. Incluso cuando ξ y η sean funciones más generales de q_r y de p_r no desarrollables en serie de potencias, las ecuaciones (9) siguen siendo suficientes para determinar el valor de $\xi\eta - \eta\xi$ como veremos más adelante. Por tanto, las ecuaciones (9) son la solución al problema de hallar las condiciones cuánticas para todos los sistemas que tengan una imagen clásica y que se puedan describir mediante variables canónicas. Sin embargo, esto no abarca todos los sistemas, de la mecánica cuántica. Las ecuaciones (7) y (9) constituyen la base de la analogía entre la mecánica cuántica y la mecánica clásica. En ellas vemos que la *mecánica clásica*

puede ser considerada como un caso límite de la cuántica cuando $\hbar$ tiende a cero. Un P.B. en mecánica cuántica es una noción puramente algebraica y, en consecuencia, es un concepto más importante que en la teoría clásica, donde sólo se puede definir en relación a un conjunto de coordenadas canónicas. Por este motivo las coordenadas canónicas son menos importantes en mecánica cuántica que en la clásica; de hecho podemos tener un sistema cuántico para el que no existan coordenadas canónicas y en el que, sin embargo, sigan teniendo sentido los P.B. Tal sistema no tendría imagen clásica, y por lo tanto no podríamos obtener sus condiciones cuánticas por el método que hemos indicado.

Según las ecuaciones (9) vemos que todo par de variables con distintos subíndices r y s conmutan siempre. De aquí se sigue que cualquier función de q_r y p_r conmuta con toda función de q_s y p_s siempre que s sea distinto de r. Distintos valores de r corresponden a distintos grados de libertad del sistema, luego *las variables dinámicas que hacen referencia a grados de libertad distintos conmutan.* Esta ley la hemos deducido de (9), y por tanto sólo la hemos demostrado para aquellos sistemas que tengan imagen clásica, pero supondremos que es válida en general. Así ya tenemos un procedimiento para hallar las primeras condiciones cuánticas de sistemas para los que, aunque no existan coordenadas canónicas, tenga sentido hablar de distintos grados de libertad, cosa que podremos hacer con la interpretación física del problema.

Ahora podemos ver cuál es el significado físico de la división de las variables dinámicas estudiada en la sección anterior, en conjuntos tales que cualquier elemento de un conjunto conmuta con cualquier elemento de los otros conjuntos. A la luz de lo dicho, cada conjunto corresponde a un cierto número de grados de libertad, o quizás a uno sólo. Dicha división puede coincidir con el proceso físico de separar el sistema dinámico en sus partes constituyentes, cada una de las cuales puede tener existencia por sí sola como un sistema físico aislado, y que puestas todas ellas en interacción dan lugar al sistema original. Pero puede ser un simple artificio matemático para separar el sistema dinámico en grados de libertad que físicamente no se pueden separar; por ejemplo, un sistema constituido por una partícula con estructura interna se puede separar en los grados de libertad que describen el movimiento del centro de la partícula y los que describen la estructura interna.

22. *Representación de Schrödinger*

Consideremos un sistema dinámico con n grados de libertad que tenga una imagen clásica, y que por lo tanto se pueda describir mediante variables canónicas q_r, p_r ($r = 1, 2, \ldots, n$). Supondremos que *las coordenadas q_r son todas observables y que tienen un espectro continuo de autovalores,*

lo que está justificado teniendo en cuenta el significado físico de la q_r. Sea una representación en la que las q_r sean diagonales. El problema consiste en saber si las q_r constituyen un conjunto completo de observables que conmutan para dicho sistema. Parece bastante natural que así sea. Por ahora lo supondremos así, y lo justificaremos más tarde (véase el final de la pág. 102). La representación en la que las q_r constituyen un conjunto completo de observables que conmutan queda completamente determinada a excepción de los factores de fase arbitrarios.

Consideremos en primer lugar el caso $n = 1$ para el que no exista más que una q y una p. Tendremos

$$qp - pq = i\hbar. \tag{10}$$

Todo ket se puede escribir en la representación del ket standard como $\psi(q)\rangle$. A partir de él podemos construir otro ket $d\psi/dq\rangle$, cuyo representante es igual a la derivada del representante del ket de partida. El ket así obtenido es función lineal del ket de partida y por tanto es el resultado de aplicar un operador lineal a dicho ket. Llamando d/dq a dicho operador, tenemos

$$\frac{d}{dq}\psi\rangle = \frac{d\psi}{dq}\rangle. \tag{11}$$

La ecuación (11), válida para toda función ψ define el operador d/dq. Tenemos

$$\frac{d}{dq}\rangle = 0. \tag{12}$$

Consideremos el operador lineal d/dq según la teoría general de operadores lineales dada en § 7. Según ella podemos aplicarlo a un bra $\langle\phi(q)$, y su producto $\langle\phi d/dq$ definido en (3) de § 7 vale

$$\left\{\langle\phi \frac{d}{dq}\right\}\psi\rangle = \langle\phi\left\{\frac{d}{dq}\psi\rangle\right\} \tag{13}$$

para toda función $\psi(q)$. Tomando representantes resulta

$$\int \langle\phi \frac{d}{dq}|q'\rangle\, dq'\, \psi(q') = \int \phi(q')\, dq' \frac{d\psi(q')}{dq'}. \tag{14}$$

Podemos transformar el segundo miembro integrando por partes, y así obtenemos

$$\int \langle\phi \frac{d}{dq}|q'\rangle\, dq'\, \psi(q') = -\int \frac{d\phi(q')}{dq'}\, dq'\, \psi(q'), \tag{15}$$

si la contribución en los límites de integración es nula. Resulta pues

$$\langle\phi \frac{d}{dq} |q'\rangle = -\frac{d\phi(q')}{dq'},$$

lo que demuestra que

$$\langle\phi \frac{d}{dq} = -\langle \frac{d\phi}{dq}. \tag{16}$$

Luego, d/dq aplicado por la izquierda a la función compleja conjugada de una función de onda representa derivar respecto a q y cambiar de signo.

La validez de esta conclusión depende de la posibilidad de que sea posible pasar de (14) a (15), y, por tanto, debemos limitarnos a considerar bras y kets que correspondan a funciones de onda que satisfagan unas condiciones de contorno adecuadas. La condición que se aplica en la práctica es que se anulen en los límites. (En la próxima sección se darán unas condiciones algo más generales.) Tales condiciones no sólo no limitan la aplicabilidad de la teoría sino que por el contrario son también necesarias frecuentemente desde el punto de vista físico. Por ejemplo, si q es una coordenada cartesiana de una partícula, sus autovalores van desde $-\infty$ a ∞, y si exigimos que la partícula tenga probabilidad nula de estar en el infinito llegamos a la condición de que la función de onda debe anularse para $q = \pm\infty$.

Podemos calcular el complejo conjugado del operador lineal d/dq teniendo en cuenta que el imaginario conjugado de $d/dq \cdot \psi\rangle$ o $d\psi/dq\rangle$ es $\langle d\psi/dq$, o sea $-\langle\psi d/dq$ según lo visto en (16). Luego, el complejo conjugado de d/dq es $-d/dq$ y, por tanto, *d/dq es un operador lineal imaginario puro.*

Para calcular el representante de d/dq tengamos presente que aplicando la fórmula (63) de § 20, se tiene

$$|q''\rangle = \delta(q - q'')\rangle, \tag{17}$$

de modo que

$$\frac{d}{dq} |q''\rangle = \frac{d}{dq} \delta(q - q'')\rangle, \tag{18}$$

y por tanto,

$$\langle q'| \frac{d}{dq} |q''\rangle = \frac{d}{dq'} \delta(q' - q''). \tag{19}$$

El representante de d/dq lleva consigo la derivada de la función δ.

Vamos a calcular la relación de conmutación entre d/dq y q. Tenemos

$$\frac{d}{dq} q\psi\rangle = \frac{dq\psi}{dq}\rangle = q \frac{d}{dq} \psi\rangle + \psi\rangle. \tag{20}$$

Puesto que esta ecuación es válida para todo ket $\psi\rangle$, resulta

$$\frac{d}{dq}q - q\frac{d}{dq} = 1. \tag{21}$$

Comparando este resultado con (10) vemos que *la relación de conmutación entre* $-i\hbar\, d/dq$ *y* q *es la misma que entre* p *y* q.

Para generalizar estos resultados al caso de un n cualquiera, designamos el ket general $\psi(q_1 \ldots q_n)\rangle = \psi\rangle$ e introducimos n operadores lineales $\partial/\partial q_r$ ($r = 1, \ldots, n$), que pueden actuar sobre dicho ket según la regla

$$\frac{\partial}{\partial q_r}\psi\rangle = \frac{\partial\psi}{\partial q_r}\rangle, \tag{22}$$

que corresponde a (11). La ecuación análoga a (12) es

$$\frac{\partial}{\partial q_r}\rangle = 0 \tag{23}$$

Si nos limitamos a considerar bras y kets que correspondan a funciones de onda que satisfacen condiciones de contorno adecuadas, estos operadores lineales pueden actuar también sobre bras según la regla

$$\langle\phi\frac{\partial}{\partial q_r} = -\langle\frac{\partial\phi}{\partial q_r}, \tag{24}$$

que corresponde a (16). Por tanto, $\partial/\partial q_r$ se puede aplicar por la izquierda a la función compleja conjugada de una función de onda con el significado de derivación respecto de q_r más cambio de signo. Igual que antes, resulta que cada uno de los operadores $\partial/\partial q_r$ es un operador lineal imaginario puro. En correspondencia con (21) resultan las relaciones de conmutación

$$\frac{\partial}{\partial q_r}q_s - q_s\frac{\partial}{\partial q_r} = \delta_{rs}. \tag{25}$$

Tenemos además

$$\frac{\partial}{\partial q_r}\frac{\partial}{\partial q_s}\psi\rangle = \frac{\partial^2\psi}{\partial q_r\partial q_s}\rangle = \frac{\partial}{\partial q_s}\frac{\partial}{\partial q_r}\psi\rangle, \tag{26}$$

de donde se deduce

$$\frac{\partial}{\partial q_r}\frac{\partial}{\partial q_s} = \frac{\partial}{\partial q_s}\frac{\partial}{\partial q_r}. \tag{27}$$

Comparando (25) y (27) con (9), vemos que *las relaciones de conmutación entre los operadores lineales* $-i\hbar\,\partial/\partial q_r$ *y los* q_s *son las mismas que entre los* p_r *y los* q_s.

Podríamos tomar

$$p_r = -i\hbar\,\partial/\partial q_r \tag{28}$$

sin dar lugar a ninguna contradicción. Tal posibilidad nos muestra que las q_r constituyen un conjunto completo de observables que conmutan, pues permite expresar toda función de las q_r y de las p_r como función de las q_r y de las $-i\hbar\, \partial/\partial q_r$ y dicha función no conmutaría con todas las q_i a menos que fuera función únicamente de las q_r.

Las ecuaciones (28) no tienen que ser válidas necesariamente. Pero, en cualquier caso, cada una de las cantidades $p_r + i\hbar\, \partial/\partial q_r$ conmuta con todas las q_s y, por tanto, según el teorema 2 de § 19, es una función de las q_s. Luego,

$$p_r = -i\hbar\, \partial/\partial q_r + f_r(q). \tag{29}$$

Como p_r y $-i\hbar\partial/\partial q_r$ son reales, $f_r(q)$ también ha de serlo. Para toda función f de las q tenemos

$$\frac{\partial}{\partial q_r} f|\psi\rangle = f\frac{\partial}{\partial q_r}|\psi\rangle + \frac{\partial f}{\partial q_r}|\psi\rangle,$$

o sea,

$$\frac{\partial}{\partial q_r} f - f\frac{\partial}{\partial q_r} = \frac{\partial f}{\partial q_r}. \tag{30}$$

Con ayuda de (29) podemos deducir la fórmula general

$$p_r f - f p_r = -i\hbar\, \partial f/\partial q_r. \tag{31}$$

Esta ecuación se puede escribir en notación de P.B.

$$[f, p_r] = \partial f/\partial q_r, \tag{32}$$

que es la misma ecuación que se deduce en la teoría clásica a partir de (1). Multiplicando (27) por $(-i\hbar)^2$ y sustituyendo $-i\hbar\, \partial/\partial q_r$ y $-i\hbar\, \partial/\partial q_s$ por sus valores dados por (29), resulta

$$(p_r - f_r)(p_s - f_s) = (p_s - f_s)(p_r - f_r),$$

que con ayuda de la condición cuántica $p_r p_s = p_s p_r$ se reduce a

$$p_r f_s + f_r p_s = = p_s f_r + f_s p_r.$$

Ésta puede a su vez simplificarse mediante la (31)

$$\partial f_s/\partial q_r = \partial f_r/\partial q_s, \tag{33}$$

que nos dice que todas las funciones f_r son de la forma

$$f_r = \partial F/\partial q_r \tag{34}$$

donde F es independiente de r. Así la ecuación (29) queda en la forma

$$p_r = -i\hbar\, \partial/\partial q_r + \partial F/\partial q_r. \tag{35}$$

Hasta aquí hemos estado trabajando en una representación que fijamos

mediante la condición de que en ella las q_r sean diagonales, pero que podía contener factores de fase arbitrarios. Si cambiamos los factores de fase cambiarán también los operadores $\partial/\partial q_r$. Vamos a demostrar ahora que eligiendo convenientemente los factores de fase puede hacerse nula la función F de (35), con lo cual serán válidas las ecuaciones (28).

Emplearemos asteriscos para señalar las cantidades que están expresadas en la nueva representación con los nuevos factores de fase. Los nuevos bras básicos estarán relacionados con los anteriores por

$$\langle q'_1 \ldots q'^*_n | = e^{i\gamma'} \langle q' \ldots q'_n | \tag{36}$$

donde $\gamma' = \gamma(q')$ es una función real de las q'. El representante de un ket en la nueva representanción será igual al anterior multiplicado por $e^{i\gamma'}$, y por tanto $e^{i\gamma}\psi\rangle^* = \psi\rangle$, de donde se deduce

$$\rangle^* = e^{-i\gamma}\rangle \tag{37}$$

que es la relación entre el ket standard de la nueva representación y el de la de partida. El nuevo operador lineal $(\partial/\partial q_r)^*$ verificará la ecuación análoga a la (22)

$$\left(\frac{\partial}{\partial q_r}\right)^* \psi\rangle^* = \frac{\partial\psi}{\partial q_r}\rangle^* = e^{-i\gamma}\frac{\partial\psi}{\partial q_r}\rangle$$

habiendo empleado (37) para la segunda igualdad. De ella, con ayuda de (22) resulta

$$\left(\frac{\partial}{\partial q_r}\right)^* \psi\rangle^* = e^{-i\gamma}\frac{\partial}{\partial q_r}\psi\rangle = e^{-i\gamma}\frac{\partial}{\partial q_r}\, e^{i\gamma}\psi\rangle^*,$$

luego,

$$\left(\frac{\partial}{\partial q_r}\right)^* = e^{-i\gamma}\frac{\partial}{\partial q_r}\, e^{i\gamma}, \tag{38}$$

y mediante (30)

$$\left(\frac{\partial}{\partial q_r}\right)^* = \frac{\partial}{\partial q_r} + i\frac{\partial\gamma}{\partial q_r}. \tag{39}$$

Si elegimos γ de modo que

$$F = \hbar\gamma + \text{constante}, \tag{40}$$

la (35) se reduce a

$$p_r = -i\hbar(\partial/\partial q_r)^*. \tag{41}$$

La ecuación (40) nos determina γ a menos de una constante de fase arbitraria, con lo que la representación queda así determinada a menos de un factor de fase constante.

En consecuencia, es posible elegir una representación donde las q_r sean diagonales y en la que sean válidas las ecuaciones (28). Esta representación

tiene gran utilidad para muchos problemas. La llamaremos *representación de Schrödinger,* pues fue la representación que empleó Schrödinger cuando en 1926 dio su formulación original de la mecánica cuántica. La representación de Schrödinger existe siempre que existan coordenadas canónicas q_r y p_r, y está perfectamente determinada a menos de un factor de fase arbitrario constante. Su gran ventaja consiste en que nos permite expresar inmediatamente toda función algebraica de las q_r y de las p_r que sea desarrollable en serie de potencias respecto a las p_r como un operador de derivación, es decir, si $f(q_1, \ldots, q_n, p_1, \ldots, p_n)$ es la función en cuestión, podemos escribir

$$f(q_1, \ldots, q_n, p_1, \ldots, p_n) = f(q_1, \ldots, q_n, -i\hbar\, \partial/\partial q_1, \ldots, -i\hbar\, \partial/\partial q_n), \quad (42)$$

conservando el orden de los factores en los productos al sustituir las p_r por $-i\hbar\partial/\partial q_r$.

De (23) y (28) resulta

$$p_r\rangle = 0. \quad (43)$$

Luego el ket standard de la representación de Schrödinger se caracteriza por ser un autoket común a todos los momentos perteneciente a los autovalores cero. Podemos considerar otras propiedades de los vectores básicos de la representación de Schrödinger. La ecuación (22) da:

$$\langle q'_1 \ldots q'_n | \frac{\partial}{\partial q_r} \psi\rangle = \langle q'_1 \ldots q'_n | \frac{\partial \psi}{\partial q'_r}\rangle = \frac{\partial \psi(q'_1 \ldots q'_n)}{\partial q'_r} =$$

$$= \frac{\partial}{\partial q'_r} \langle q'_1 \ldots q'_n | \psi\rangle.$$

Luego,

$$\langle q'_1 \ldots q'_n | \frac{\partial}{\partial q_r} = \frac{\partial}{\partial q'_r} \langle q'_1 \ldots q'_n |, \quad (44)$$

y por tanto,

$$\langle q'_1 \ldots q'_n | p_r = -i\hbar \frac{\partial}{\partial q'_r} \langle q'_1 \ldots q'_n |. \quad (45)$$

Análogamente, la ecuación (24) nos lleva a

$$p_r | q'_1 \ldots q'_n\rangle = i\hbar \frac{\partial}{\partial q'_r} | q'_1 \ldots q'_n\rangle. \quad (46)$$

23. *La representación de momentos*

Sea un sistema de un sólo grado de libertad, que se pueda describir mediante una q y una p siendo el espectro de autovalores de la q el conjunto de todos los números desde $-\infty$ hasta $+\infty$, y consideremos un autoket $|p'\rangle$ de p. Su representante $\langle q'|p'\rangle$ en la representación de Schrödinger verificará

$$p'\langle q'|p'\rangle = \langle q'|p|p'\rangle = -i\hbar \frac{d}{dq'} \langle q'|p'\rangle,$$

como resulta de aplicar (45) para el caso de un sólo grado de libertad. La solución de esta ecuación diferencial en $\langle q'|p'\rangle$ es

$$\langle q'|p'\rangle = c' \, e^{ip'q'/\hbar}, \tag{47}$$

siendo $c' = c(p')$ independiente de q', pero pudiendo depender de p'.

El representante $\langle q'|p'\rangle$ no verifica la condición de contorno de anularse para $q' = \pm\infty$. Esto lleva consigo ciertas dificultades, que se ponen claramente de manifiesto al no cumplirse el teorema de ortogonalidad. Si consideramos otro autoket $|p''\rangle$ de p de representante

$$\langle q'|p''\rangle = c'' \, e^{ip''q'/\hbar},$$

que pertenezca a un autovalor p'' distinto, resultará

$$\langle p'|p''\rangle = \int_{-\infty}^{\infty} \langle p'|q'\rangle \, dq' \, \langle q'|p''\rangle = \bar{c}'c'' \int_{-\infty}^{\infty} e^{-i(p'-p'')q'/\hbar} \, dq'. \tag{48}$$

Esta integral no converge en el sentido de la definición habitual de convergencia. Para que la teoría sea coherente, adoptamos una nueva definición de convergencia para las integrales cuyo dominio de integración se extienda a infinito, análoga a la definición de Cesàro para la suma de series infinitas. Con la nueva definición, una integral que en el límite superior se comporte como cos aq' o sen aq', siendo a un número real distinto de cero, se dice que no contribuye al valor de la integral en dicho límite, es decir, se toma el valor medio de las oscilaciones. Y análogamente para el límite inferior cuando q' tiende a menos infinito. Con esta definición, el segundo miembro de (48) se anula para $p'' \neq p'$, y el teorema de ortogonalidad sigue cumpliéndose. Asimismo son iguales los segundos miembros de (13) y (14) cuando $\langle\phi$ y $\psi\rangle$ son autovectores de p, con lo cual el operador d/dq se puede aplicar a los autovectores de p. Luego, las condiciones de contorno que han de verificar los representantes de los bras y de los kets que tienen sentido se puede generalizar exigiendo tan sólo que cuando q' tiende a más o menos infinito, el representante se comporte como cos aq' o sen aq'.

Cuando p'' se aproxima a p', el segundo miembro de (48) se convierte en una función δ. Para calcularlo necesitamos la fórmula

$$\int_{-\infty}^{\infty} e^{iax}\, dx = 2\pi\, \delta(a) \tag{49}$$

que es válida para a real, como vamos a demostrar a continuación. Para a distinto de cero los dos miembros son nulos, y la fórmula evidentemente es correcta. Además, para toda función continua $f(a)$ cuando g tiende a infinito, se tiene

$$\int_{-\infty}^{\infty} f(a)\, da \int_{-g}^{g} e^{iax}\, dx = \int_{-\infty}^{\infty} f(a)\, da\, 2a^{-1} \operatorname{sen} ag = 2\pi f(0)$$

Puede darse un argumento más complicado para demostrar que se llega al mismo resultado tomando en lugar de g y $-g$, los límites g_1 y $-g_2$ que tiendan a infinito de modo distinto (aunque no completamente distinto). Esto nos demuestra la equivalencia de los dos miembros de (49) cuando los utilizamos como factores bajo el signo integral, quedando así demostrada la fórmula.

Teniendo en cuenta (49), la (48) se convierte en

$$\begin{aligned}\langle p'|p''\rangle = \bar{c}'c''\, 2\pi\, \delta[(p'-p'')/\hbar] &= \bar{c}'c''\, h\, \delta(p'-p'') \\ &= |c'|^2\, h\, \delta(p'-p'').\end{aligned} \tag{50}$$

Hemos obtenido un autoket de p para cada autovalor real p', y su representante viene dado por (47). Todo ket $|X\rangle$ se puede expresar en función de dichos autokets de p, pues su representante $\langle q'|X\rangle$ se puede desarrollar en serie de los representantes (47) utilizando el análisis de Fourier. Por tanto, *el momento p es un observable,* de acuerdo con el resultado experimental de que los momentos se pueden observar.

Así aparece una simetría entre q y p. Ambos son observables cuyos autovalores van desde $-\infty$ a $+\infty$, y la relación de conmutación entre ellos (10), no cambia si intercambiamos q con p y sustituimos i por $-i$. Hemos establecido una representación en la que q es diagonal, y donde $p = -i\hbar\, d/dq$. De la simetría entre q y p se deduce que también podemos establecer una representación en la que p sea diagonal y donde

$$q = i\hbar\, d/dp, \tag{51}$$

El operador d/dp lo definiremos de forma análoga a como hicimos para d/dq. A esta representación se le denomina *representación de momentos.* Es menos útil que la representación de Schrödinger de la sección anterior, pues ésta nos permite expresar toda función de q y p, que sea desarrollable en serie de potencias respecto a p como un operador de derivación, mientras que la representación de momentos nos permite expresar como opera-

dor de derivación las funciones de q y p que sean desarrollables en serie respecto a q, y las cantidades más importantes de la dinámica casi siempre son series de potencias de p y muy pocas veces son series de potencias de q. Sin embargo, la representación de momentos tiene utilidad para algunos problemas (véase § 50).

Calculemos la función de transformación $\langle q'|p'\rangle$ que relaciona ambas representaciones. Los kets básicos $|p'\rangle$ de la representación de momentos son autokets de p, y sus representantes en la representación de Schrödinger $\langle q'|p'\rangle$ están dados por (47) con unos coeficientes c' convenientemente elegidos. Los factores de fase de estos kets básicos se tienen que elegir de tal modo que (51) sea válida. El método más sencillo de obtener esta condición es considerar la simetría que hemos visto antes entre q y p, según la cual $\langle q'|p'\rangle$ se ha de transformar en $\langle p'|q'\rangle$ al intercambiar q' con p' y escribir $-i$ en lugar de i. Pero $\langle q'|p'\rangle$ es igual al segundo miembro de (47) y $\langle p'|q'\rangle$ es su complejo conjugado, de donde se sigue que c' ha de ser independiente de p'. Luego, c' es un número c. Además debe verificarse

$$\langle p'|p''\rangle = \delta(p' - p''),$$

que comparada con (50) nos indica que $|c| = h^{-\frac{1}{2}}$. Podemos elegir los factores de fase constantes y arbitrarios de cada representación de modo que sea $c = h^{-\frac{1}{2}}$, con lo que resulta la siguiente función de transformación:

$$\langle q'|p'\rangle = h^{-\frac{1}{2}} e^{ip'q'/\hbar} \tag{52}$$

Todo esto puede generalizarse a un sistema con n grados de libertad, para el que existen n coordenadas q_i y p_i, y para el que los autovalores de cada una de las q_i vayan desde $-\infty$ a $+\infty$. Entonces cada p_i será un observable cuyos autovalores irán también desde $-\infty$ a ∞, y existirá una simetría entre el conjunto de q_i y el de p_i cuyas relaciones de conmutación permanecerán invariantes al intercambiar cada q_r con su correspondiente p_r y poner $-i$ en lugar de i. Se podrá establecer una representación con las p_i diagonales y en la que cada

$$q_r = i\hbar\partial/\partial p_r \tag{53}$$

La función de transformación que nos relacionará esta representación con la de Schrödinger vendrá dada por el producto de las funciones de transformación relativas a cada grado de libertad, según indica la fórmula (67) de § 20 y, por tanto, será

$$\begin{aligned}\langle q'_1 q'_2 \ldots q'_n | p'_1 p'_2 \ldots p'_n\rangle &= \langle q'_1|p'_1\rangle\langle q'_2|p'_2\rangle \ldots \langle q'_n|p'_n\rangle \\ &= h^{-n/2} e^{i(p'_1 q'_1 + p'_2 q'_2 + \cdots + p'_n q'_n)/\hbar}\end{aligned} \tag{54}$$

24. *El principio de incertidumbre de Heisenberg*

Para un sistema que no tenga más que un grado de libertad, los representantes de un ket $|X\rangle$ en las representaciones de Schrödinger y de momentos están relacionados por

$$\left.\begin{aligned} \langle p'|X\rangle &= h^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{-iq'p'/\hbar}\, dq'\, \langle q'|X\rangle, \\ \langle q'|X\rangle &= h^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{iq'p'/\hbar}\, dp'\, \langle p'|X\rangle. \end{aligned}\right\} \quad (55)$$

Estas fórmulas tienen un significado elemental. Nos dicen que *cada uno de los representantes viene dado, salvo coeficientes numéricos, por las amplitudes de las componentes de Fourier del otro.*

Es interesante aplicar (55) a un ket cuyo representante en la representación de Schrödinger sea lo que se denomina un *paquete de ondas.* Esto quiere decir, una función cuyo valor sea muy pequeño fuera de un cierto dominio de amplitud $\Delta q'$, dentro del cual sea aproximadamente periódico con una frecuencia definida.* Si hacemos el análisis de Fourier de este paquete de ondas, todas las amplitudes de las componentes de Fourier serán pequeñas excepto las del entorno de una frecuencia definida. Las componentes cuya amplitud no será pequeña estarán dentro de una banda de frecuencias * de anchura aproximada $1/\Delta q'$, pues dos componentes cuyas frecuencias difieran en esta cantidad, si están en fase en el medio del dominio $\Delta q'$, estarán en oposición de fase e interferirán hacia el extremo de dicho dominio. En la primera de las ecuaciones (55), la variable que hace las veces de frecuencia es $(2\pi)^{-1}p'/\hbar = p'/h$. Luego, si $\langle q'|X\rangle$ tiene la forma de un paquete de ondas, la función $\langle p'|X\rangle$ que viene dada por las amplitudes de Fourier del paquete de ondas tendrá un valor pequeño en todo el espacio de p' salvo en un cierto dominio de anchura $\Delta p' = h/\Delta q'$.

Apliquemos ahora la interpretación física según la cual el cuadrado del módulo del representante de un ket representa una probabilidad.

Resulta entonces que nuestro paquete de onda representa un estado en el que casi con certeza toda medida de q da un valor comprendido en un cierto dominio de anchura $\Delta q'$ y toda medida de p da un valor comprendido en un dominio de anchura $\Delta p'$. Diremos que en este estado q tiene un valor definido con un error del orden de $\Delta q'$, y que asimismo p tiene un valor definido con un error del orden de $\Delta p'$. El producto de los dos errores es

---

* Aquí, frecuencia quiere decir el recíproco de la longitud de onda. [*N. del T.*: Es decir, la palabra frecuencia se emplea aquí para designar el número de ondas.]

$$\Delta q' \Delta p' = h. \tag{56}$$

Es decir, cuanto mayor sea la precisión sobre una de las variables p o q tanto menor será la precisión sobre la otra. En un sistema con más de un grado de libertad la ecuación (56) se aplica por separado a cada uno.

La ecuación (56) se conoce con el nombre de *principio de incertidumbre de Heisenberg*. En ella se ve claramente el grado de limitación que existe en asignar simultáneaemnte valores numéricos en un estado particular a dos observables que, como las coordenadas canónicas conjugadas, no conmutan, y constituye un ejemplo claro de en qué sentido pueden ser incompatibles las medidas en mecánica cuántica. También nos indica por qué la mecánica clásica, que supone la posibilidad de asignar simultáneamente valores numéricos a todos los observables, es una aproximación válida cuando h se puede considerar suficientemente pequeña para ser despreciada. La ecuación (56) solamente es válida en el caso más favorable de que el representante del estado sea de la forma de un paquete de ondas. Para otras formas de representante obtendríamos unos $\Delta q'$ y $\Delta p'$ cuyo producto sería mayor que h.

El principio de incertidumbre de Heisenberg nos indica que en el límite, cuando una de las variables q o p esté completamente determinada, la otra estará completamente indeterminada. También podríamos haber llegado a este mismo resultado directamente a partir de la función de transformación $\langle q'|p'\rangle$. Según dijimos al final de § 18, $|\langle q'|p'\rangle|^2 dq'$ es proporcional a la probabilidad de que q tenga un valor comprendido en el pequeño intervalo $(q', q' + dq')$ en el estado en que p tiene con certeza el valor p', y según (52) dicha probabilidad es independiente de q' para un dq' dado. Luego, si p tiene con toda seguridad el valor definido p', todos los valores de q son igualmente probables. Análogamente, si q tiene con toda seguridad un valor definido q', todos los valores de p son igualmente probables.

Desde el punto de vista físico es evidente la imposibilidad de obtener un estado para el que sean igualmente probables todos los valores de q o todos los valores de p; en el primer caso debido a la limitación de espacio y en el segundo a la limitación de energía. Por lo tanto, en la práctica no se pueden realizar los autoestados de p y q. El argumento del final de § 12 ya nos indicaba que no era posible realizar dichos estados basándose en que para ello sería necesaria una precisión infinita, y ahora hemos dado otro argumento que nos conduce a la misma conclusión.

25. *Operadores de traslación*

Obtendremos nueva luz acerca del significado de algunas condiciones cuánticas tras haber estudiado los operadores de traslación. Estos operadores aparecen en la teoría al considerar que el esquema dado en el capí-

tulo II, que relaciona los estados y las variables dinámicas, es esencialmente un esquema *físico*, y consiguientemente si existe una relación entre ciertos estados y variables dinámicas, trasladándolos todos del mismo modo (por ejemplo, trasladándolos en una magnitud δx en la dirección del eje x de unas coordenadas cartesianas) los correspondientes estados y variables dinámicas trasladados tienen que seguir verificando la misma relación.

La traslación de un estado o de un observable son procesos perfectamente definidos desde el punto de vista físico. En efecto, para trasladar un estado o un observable una distancia δx en la dirección del eje x, basta con trasladar todos los aparatos utilizados para preparar el estado o todos los aparatos utilizados para medir el observable en una distancia δx en dicha dirección. Los aparatos trasladados definirán el estado o el observable trasladado. La traslación de una variable dinámica ha de estar tan bien definida como la traslación de un observable, dada la estrecha relación matemática que existe entre variables dinámicas y observables. Un estado o una variable dinámica trasladados deben estar perfectamente definidos por el estado o la variable dinámica sin trasladar y la dirección y magnitud de la traslación.

En cambio, la traslación de un ket no está tan claramente definida. Si tomamos un cierto ket, éste representa un cierto estado que por traslación se convierte en un nuevo estado perfectamente definido; pero el nuevo estado no determina unívocamente al ket trasladado, sino únicamente su dirección. Para ayudar a determinar el ket trasladado podemos exigir que tenga la misma longitud que el ket sin trasladar, pero aún así sigue sin estar completamente determinado, y se le puede multiplicar por un factor de fase arbitrario. A primera vista podríamos pensar que cada ket trasladado podría tener un factor de fase distinto, pero vamos a ver que tiene que ser el mismo para todos. Haremos uso de la regla según la cual las relaciones de superposición entre estados han de ser invariantes frente a las traslaciones. Una relación de superposición entre estados se expresa matemáticamente por una ecuación lineal entre los kets que corresponden a dichos estados, como por ejemplo

$$|R\rangle = c_1|A\rangle + c_2|B\rangle, \tag{57}$$

donde c_1 y c_2 son números; y la invariancia de la relación de superposición se expresará diciendo que los estados trasladados han de corresponder a kets que satisfagan la misma ecuación lineal de antes — en nuestro ejemplo corresponderán a unos $|Rd\rangle$, $|Ad\rangle$ y $|Bd\rangle$ que habrán de verificar

$$|Rd\rangle = c_1|Ad\rangle + c_2|Bd\rangle. \tag{58}$$

Tomaremos como kets trasladados a éstos y no a éstos multiplicados por un factor de fase independiente para cada uno de ellos, en cuyo caso verificarían una ecuación lineal con coeficientes c_1 y c_2 distintos. Ahora la única arbitrariedad en la elección de los kets trasladados es un único factor de fase arbitrario que puede multiplicar a todos.

La condición de que las ecuaciones lineales entre kets sean invariantes frente a las traslaciones, o sea, de que si es válida una ecuación cualquiera de la forma (57) lo sea también su correspondiente (58), significa que los kets trasladados son funciones lineales de los kets sin trasladar y que, en consecuencia, cualquier ket $|Pd\rangle$ es el resultado de aplicar un cierto operador lineal al correspondiente ket $|P\rangle$ no trasladado. Es decir,

$$|Pd\rangle = D|P\rangle, \tag{59}$$

siendo D un operador lineal independiente de $|P\rangle$ y que sólo depende de la traslación. El hecho de que los kets trasladados puedan ser multiplicados por un factor de fase arbitrario se traduce en que D esté indeterminado en un factor numérico arbitrario de módulo unidad.

Habiendo fijado los kets trasladados tal como hemos hecho, y en consecuencia, fijando también los bras trasladados por la condición de que sean los imaginarios conjugados de los kets, podemos afirmar que cualquier ecuación simbólica entre kets, bras y variables dinámicas tiene que permanecer invariante cuando trasladamos todos los símbolos que aparecen en ella, pues dicha ecuación tiene un significado físico que no puede cambiar por la traslación.

Tomemos como ejemplo la ecuación

$$\langle Q|P\rangle = c,$$

en la que c es un número. Entonces debe verificarse

$$\langle Qd|Pd\rangle = c = \langle Q|P\rangle. \tag{60}$$

La conjugada imaginaria de (59), poniendo Q en lugar de P da

$$\langle Qd| = \langle Q|\bar{D}. \tag{61}$$

De (60) obtenemos ahora

$$\langle Q|\bar{D}D|P\rangle = \langle Q|P\rangle.$$

y como esto es válido para cualesquiera $\langle Q|$ y $|P\rangle$, resulta

$$\bar{D}D = 1, \quad \surd \tag{62}$$

que es una condición general que debe satisfacer D.

Como segundo ejemplo consideremos la ecuación.

$$v|P\rangle = |R\rangle,$$

siendo v una variable dinámica cualquiera. Entonces, si v_d representa la variable dinámica trasladada, debe verificarse

$$v_d|Pd| = |Rd\rangle.$$

Con ayuda de (59) obtenemos

$$v_d|Pd\rangle = D|R\rangle = Dv|P\rangle = DvD^{-1}|Pd\rangle.$$

y como $|Pd\rangle$ es un ket cualquiera, debe ser

$$v_d = DvD^{-1}, \tag{63}$$

ecuación que indica que el operador lineal D no sólo determina la traslación de kets y bras, sino también de las variables dinámicas. Nótese que el factor numérico arbitrario de módulo unidad que puede multiplicar a D no afecta a v_d, y por tanto, no afecta tampoco la validez de (62).

Pasemos a considerar ahora una traslación infinitesimal, es decir, consideremos una traslación δx en la dirección del eje x y hagamos que $\delta x \to 0$. Por continuidad física hemos de suponer que el ket trasladado $|Pd\rangle$ tiende al ket original $|P\rangle$, y asimismo debemos suponer que existe el límite

$$\lim_{\delta x \to 0} \frac{|Pd\rangle - |P\rangle}{\delta x} = \lim_{\delta x \to 0} \frac{D-1}{\delta x}|P\rangle$$

Ello exige que exista el límite

$$\lim_{\delta x \to 0} (D-1)/\delta x \tag{64}$$

Tal límite es un operador lineal que llamaremos el *operador de traslación* en la dirección del eje x, y lo representaremos por d_x. El factor numérico $e^{i\gamma}$ con γ real que puede multiplicar a D tiene que tender a la unidad cuando $\delta x \to 0$, lo que introduce una arbitrariedad en d_x, que puede ser sustituido por

$$\lim_{\delta x \to 0} (De^{i\gamma} - 1)/\delta x = \lim_{\delta x \to 0} (D - 1 + i\gamma)/\delta x = d_x + ia_x,$$

siendo a_x el límite de $\gamma/\delta x$. Por tanto, d_x contiene un término aditivo constituido por un número imaginario puro arbitrario.

Para pequeños δx,

$$D = 1 + \delta x\, d_x. \tag{65}$$

Sustituyendo esto en (62) resulta

$$(1 + \delta x\, d_x)(1 + \delta x\, d_x) = 1,$$

que despreciando δx^2 se reduce a

$$\delta x(\bar{d}_x + d_x) = 0.$$

Luego, d_x es un operador lineal imaginario puro. Sustituyendo (65) en (63) y despreciando de nuevo δx^2 resulta

$$v_d = (1 + \delta x\, d_x)v(1 - \delta x\, d_x) = v + \delta x(d_x\, v - v\, d_x), \tag{66}$$

que demuestra que

$$\lim_{\delta x \to 0} (v_d - v)/\delta x = d_x\, v - v\, d_x. \tag{67}$$

Para describir un sistema dinámico cualquiera podemos utilizar las siguientes variables dinámicas: las coordenadas cartesianas x, y, z del centro de gravedad del sistema, las componentes p_x, p_y, p_z del momento total del sistema, que son las coordenadas canónicas conjugadas respectivas de x, y, z, y cualesquiera otras variables dinámicas que sean necesarias para describir los grados de libertad internos. Si el conjunto de aparatos que nos servían para medir x lo trasladamos una distancia δx en la dirección del eje x, medirá $x - \delta x$, luego,

$$x_d = x - \delta x.$$

Comparando este resultado con (66) para $v = x$ obtenemos

$$d_x x - x d_x = -1, \tag{68}$$

que es la relación de conmutación entre d_x y x. Con argumentos análogos veríamos que y, z, p_x, p_y, p_z, así como las variables dinámicas internas, al no estar afectadas por la traslación, tienen que conmutar con d_x. Comparando estos resultados con (9) vemos que $i\hbar d_x$ verifica exactamente las mismas relaciones de conmutación que p_x. Su diferencia, $p_x - i\hbar d_x$, conmuta con todas las variables dinámicas, y por tanto tiene que ser un número. Dicho número, que es real necesariamente, pues p_x y $i\hbar d_x$ son ambos reales, puede hacerse igual a cero eligiendo convenientemente el número imaginario puro que se podía sumar a d_x. En ese caso resulta

$$p_x = i\hbar\, d_x, \tag{69}$$

o sea, que *la componente x del momento lineal total del sistema es igual al operador de traslación d_x multiplicado por $i\hbar$.*

Este es un resultado fundamental que da un nuevo significado a los operadores de traslación. Análogamente llegamos a conclusiones correspondientes para y y z con los operadores de traslación d_y y d_z. Las condiciones cuánticas que nos indican que p_x, p_y y p_z conmutan entre sí pueden interpretarse relacionándolas con el hecho de que las traslaciones según direcciones distintas son operaciones conmutables.

26. *Transformaciones unitarias*

Sea U un operador lineal cualquiera que tenga recíproco U^{-1} y consideremos la ecuación

$$\alpha^* = U\alpha U^{-1}, \tag{70}$$

en la que α es un operador lineal cualquiera. Esta ecuación puede considerarse como una transformación que a cada operador lineal α le hace corresponder un operador lineal α^*, y en este sentido tiene propiedades muy notables. En primer lugar, todo α^* tiene los mismos autovalores que

su correspondiente α; pues si α' es un autovalor de α y $|\alpha'\rangle$ es un autoket perteneciente a él, se tiene

$$\alpha|\alpha'\rangle = \alpha'|\alpha'\rangle$$

y, por tanto,

$$\alpha^* U|\alpha'\rangle = U\alpha U^{-1}U|\alpha'\rangle = U\alpha|\alpha'\rangle = \alpha' U|\alpha'\rangle,$$

que demuestra que $U|\alpha'\rangle$ es un autoket de α^* perteneciente al mismo autovalor α'; y análogamente podemos ver que todo autovalor de α^* es autovalor de α. Además, si consideramos diversos α que estén relacionados por ecuaciones algebraicas, y los transformamos de acuerdo con (70), los correspondientes α^* estarán ligados por las mismas ecuaciones algebraicas. Esta conclusión se basa en el hecho de que los procesos algebraicos fundamentales de suma y multiplicación son invariantes frente a la transformación (70), según indican las siguientes ecuaciones:

$$(\alpha_1 + \alpha_2)^* = U(\alpha_1 + \alpha_2)U^{-1} = U\alpha_1 U^{-1} + U\alpha_2 U^{-1} = \alpha_1^* + \alpha_2^*,$$
$$(\alpha_1 \alpha_2)^* = U\alpha_1 \alpha_2 U^{-1} = U\alpha_1 U^{-1} U\alpha_2 U^{-1} = \alpha_1^* \alpha_2^*$$

Veamos cuál es la condición que deberíamos imponer a U para que transformase todo operador α real en otro α^* también real. La ecuación (70) se puede escribir

$$\alpha^* U = U\alpha. \tag{71}$$

Tomando el complejo conjugado en ambos miembros, de acuerdo con (5) de § 8, si α y α^* son reales, se tiene

$$\bar{U}\alpha^* = \alpha\bar{U}. \tag{72}$$

La ecuación (71) da

$$\bar{U}\alpha^* U = \bar{U}U\alpha$$

y la (72) da

$$\bar{U}\alpha^* U = \alpha\bar{U}U.$$

Luego,

$$\bar{U}U\alpha = \alpha\bar{U}U.$$

Por tanto, $\bar{U}U$ conmuta con cualquier operador lineal real y en consecuencia conmuta también con cualquier operador lineal, pues todo operador lineal se puede descomponer en uno real y otro también real multiplicado por i. Luego, $\bar{U}U$ es un número. Evidentemente es real, pues de acuerdo con (5) de § 8, su complejo conjugado es igual a él; y además tiene que ser un número positivo, ya que para cualquier ket $|P\rangle$, $\langle P|\bar{U}U|P\rangle$ es positivo, igual que $\langle P|P\rangle$. Podemos suponer que es igual a uno sin pérdida de generalidad en la transformación (70). Entonces tenemos

$$\bar{U}U = 1. \tag{73}$$

La ecuación (73) es equivalente a cualquiera de éstas:

$$U = \bar{U}^{-1}, \bar{U} = U^{-1}, U^{-1}\bar{U}^{-1} = 1. \tag{74}$$

Una matriz o un operador lineal U que verifique (73) y (74) se dice que es *unitario*, y una transformación (70) con U unitario se dice que es una *transformación unitaria*. Una transformación unitaria transforma operadores lineales reales en operadores lineales reales y deja invariante cualquier ecuación algebraica entre operadores lineales. Puede considerarse que actúa sobre kets y bras según las siguientes ecuaciones

$$|P^*\rangle = U|P\rangle, \qquad \langle P^*| = \langle P|\bar{U} = \langle P|U^{-1}, \tag{75}$$

y entonces deja invariante también cualquier ecuación algebraica entre operadores lineales, kets y bras. Transforma autovectores de α en autovectores de α^*. De aquí se deduce fácilmente que transforma observables en observables y que deja invariante toda relación funcional entre observables basada en la definición general de función dada en § 11.

La inversa de una transformación unitaria es a su vez una transformación unitaria, pues según (74), si U es unitario, U^{-1} también lo es. Además, si realizamos dos transformaciones unitarias una detrás de otra, el resultado es una tercera transformación unitaria, como vamos a ver seguidamente. Sean (70) y

$$\alpha^\dagger = V\alpha^* V^{-1}.$$

las dos transformaciones unitarias. La relación entre $\alpha^\dagger$ y α, según (42) de § 11 será entonces

$$\begin{aligned} \alpha^\dagger &= VU\alpha U^{-1}V^{-1} \\ &= (VU)\alpha(VU)^{-1} \end{aligned} \tag{76}$$

Pero VU es unitaria, ya que

$$\bar{V}\bar{U}VU = \bar{U}\bar{V}VU = \bar{U}U = 1,$$

luego, (76) es una transformación unitaria.

Las transformaciones de traslación dadas en la sección anterior constituyen un ejemplo de transformaciones unitarias, pues como se ve, las ecuaciones (62) y (63) corresponden a (73) y (70), y las (59) y (61) corresponden a las ecuaciones (75).

En mecánica clásica puede hacerse un cambio de coordenadas canónicas y pasar de q_r, p_r ($r = 1, \ldots, n$) a un nuevo conjunto de variables q_r^*, p_r^* ($r = 1, \ldots, n$), que tienen los mismos P.B. entre ellas que las q_r y p_r, es decir, que verifican las ecuaciones (8) de § 21 poniendo q_r^* y p_r^* en lugar de las q_r y p_r, y de tal modo que toda variable dinámica se puede expresar como función de las q_r^* y p_r^*. Se dice que las q_r^* y p_r^* constituyen otro conjunto de coordenadas canónicas, y el cambio de coordenadas se llama *transformación de contacto*. Puede comprobarse fácilmente que el P.B. de las variables u y v viene dado en función de las q_r^* y p_r^* por la misma fórmula (1) de § 21 sustituyendo q_r y p_r por q_r^* y p_r^*, luego, los P.B. son invariantes frente a las transformaciones de contacto. Esto se traduce

en que las nuevas coordenadas canónicas están en pie de igualdad con las antiguas para muchos fines de la teoría dinámica general, si bien q_r^* y p_r^* puede no ser un conjunto de coordenadas lagrangianas, sino ser funciones de las coordenadas y de las velocidades lagrangianas.

Vamos a demostrar ahora que para un sistema dinámico de la mecánica cuántica que tenga una imagen clásica, *las transformaciones unitarias de la teoría cuántica corresponden a las transformaciones de contacto de la teoría clásica.* Las transformaciones unitarias son más generales que las transformaciones de contacto, pues las primeras se pueden aplicar a sistemas cuánticos que no tengan imagen clásica, pero para los sistemas de la mecánica cuántica que se pueden describir en función de coordenadas canónicas es válida la analogía. Para establecerla veamos que si aplicamos una transformación unitaria a unas variables cuántics q_r y p_r, obtenemos unas nuevas variables q_r^* y p_r^* que tienen los mismos P.B., pues los P.B. son equivalentes a las ecuaciones algebraicas (9) de § 21, y las ecuaciones algebriacas son invariantes frente a las transformaciones unitarias. Recíprocamente, todo conjunto de variables reales q_r^* y p_r^* que tengan unos P.B. iguales que los correspondientes a coordenadas canónicas están ligados a las q_r y p_r por una transformación unitaria, como vamos a ver a continuación.

Utilizaremos la representación de Schrödinger, y para abreviar escribiremos los kets básicos $|q'_1 \ldots q'_n\rangle$ simplemente $|q'\rangle$. Como hemos supuesto que las q_r^* y p_r^* tienen unos P.B. iguales a los que corresponden a un conjunto de coordenadas canónicas, podemos establecer para ellos una representación de Schrödinger en la que cada q_r^* sea diagonal y en la que cada p_r^* sea igual a $-i\hbar\partial/\partial q_r^*$. Los kets básicos de esta segunda representación de Schrödinger serán $|q_1^{*\prime} \ldots q_n^{*\prime}\rangle$, que para abreviar designaremos $|q^{*\prime}\rangle$. Introduzcamos el operador lineal U definido mediante la relación

$$\langle q^{*\prime}|U|q'\rangle = \delta(q^{*\prime} - q'), \tag{77}$$

donde para abreviar, $\delta(q^{*\prime} - q')$ indica

$$\delta(q^{*\prime} - q') = \delta(q_1^{*\prime} - q'_1)\delta(q_2^{*\prime} - q'_2)\ldots\delta(q_n^{*\prime} - q'_n). \tag{78}$$

La ecuación compleja conjugada de (77) es

$$\langle q'|\bar{U}|q^{*\prime}\rangle = \delta(q^{*\prime} - q'),$$

luego,†

---

† Utilizamos la notación abreviada de escribir un solo signo integral y $dq^{*\prime}$ para indicar una integración respecto a las variables $q_1^{*\prime}$, $q_2^{*\prime}$, ..., $q_n^{*\prime}$. Esta abreviación se seguirá utilizando de aquí en adelante.

$$\langle q'|\bar{U}U|q''\rangle = \int \langle q'|\bar{U}|q^{*\prime}\rangle\, dq^{*\prime}\, \langle q^{*\prime}|U|q''\rangle$$
$$= \int \delta(q^{*\prime} - q')\, dq^{*\prime}\, \delta(q^{*\prime} - q'')$$
$$= \delta(q' - q''),$$

y por tanto,

$$\bar{U}U = 1.$$

En consecuencia, U es un operador unitario. Tenemos además que

$$\langle q^{*\prime}|q_r^* U|q'\rangle = q_r^{*\prime}\delta(q^{*\prime} - q')$$

y

$$\langle q^{*\prime}|Uq_r|q'\rangle = \delta(q^{*\prime} - q')q_r'.$$

Los segundos miembros de ambas ecuaciones son iguales según la propiedad (11) de § 15 de la función δ, y por tanto,

$$q_r^* U = Uq_r$$

o bien

$$q_r^* = Uq_r\, U^{-1}.$$

Además, de (45) y (46) resulta

$$\langle q^{*\prime}|p_r^* U|q'\rangle = -i\hbar \frac{\partial}{\partial q_r^{*\prime}} \delta(q^{*\prime} - q'),$$

$$\langle q^{*\prime}|Up_r|q'\rangle = i\hbar \frac{\partial}{\partial q_r'} \delta(q^{*\prime} - q').$$

Los segundos miembros de estas dos ecuaciones son evidentemente iguales, y, por tanto,

$$p_r^* U = Up_r$$

o bien

$$p_r^* = Up_r\, U^{-1}.$$

Luego, se verifican todas las condiciones que se exigen para que la transformación sea unitaria.

Para obtener una transformación unitaria infinitesimal, tomemos U en (70) de forma que difiera de la unidad en un infinitésimo. Pongamos

$$U = 1 + i\varepsilon F,$$

donde ε es un infinitésimo, y su cuadrado puede ser despreciado. Entonces

$$U^{-1} = 1 - i\varepsilon F.$$

La condición (73) o (74) de que U sea unitario exige que F sea real. La ecuación de transformación (70) se traduce en

$$\alpha^* = (1 + i\varepsilon F)\alpha(1 - i\varepsilon F),$$

que nos da

$$\alpha^* - \alpha = i\varepsilon(F\alpha - \alpha F). \tag{79}$$

Esta ecuación se puede escribir en notación de P.B.

$$\alpha^* - \alpha = \varepsilon\hbar[\alpha, F]. \tag{80}$$

Si α es una coordenada canónica, esta ecuación equivale formalmente a una transformación de contacto infinitesimal de la mecánica clásica.

V

LAS ECUACIONES DEL MOVIMIENTO

27. *Ecuaciones del movimiento en imagen de Schrödinger*

Todo nuestro desarrollo desde el § 5 se ha limitado a considerar un solo instante de tiempo. Hemos expuesto el esquema general de las relaciones entre estados y variables dinámicas para un sistema dinámico en un instante de tiempo. Para completar la teoría dinámica, hemos de considerar a su vez la conexión entre diferentes instantes de tiempo. Cuando se lleva a cabo una observación del sistema dinámico, su estado cambia de forma impredictible, pero entre observación y observación sigue habiendo causalidad tanto en mecánica cuántica como en la clásica, y el sistema está regido por ecuaciones de movimiento que hacen que el estado en un cierto instante determine el estado en un instante posterior. Vamos a estudiar a continuación estas ecuaciones del movimiento. Se aplicarán mientras el sistema no sea alterado por una observación o un proceso análogo.† Su expresión general se puede deducir del principio de superposición del capítulo I.

Consideremos un estado de movimiento particular durante el tiempo en que el sistema permanece inalterado. En cada instante t el estado corresponderá a un cierto ket que, en consecuencia, dependerá de t y al que podemos designar $|t\rangle$. Cuando consideremos varios estados de movimiento, los distinguiremos unos de otros añadiéndoles otros símbolos, tales como A, y así escribiremos el ket que corresponde a uno de estos estados en el tiempo t, $|At\rangle$. La exigencia de que el estado en un instante determine el estado en un instante posterior, quiere decir que $|At_0\rangle$ determina $|At\rangle$ a menos de un factor numérico. El principio de superposición se aplica a dichos estados de movimiento durante el tiempo en que el sistema permanece inalterado, y dice que si tomamos una relación de superposición válida

† La preparación de un estado es un proceso de este tipo. A menudo consiste en llevar a cabo una observación y seleccionar el sistema cuando el resultado de la observación es con seguridad un cierto número prefijado.

para ciertos estados en el tiempo t_0 y que da lugar a una ecuación lineal entre los correspondientes kets, como por ejemplo

$$|Rt_0\rangle = c_1|At_0\rangle + c_2|Bt_0\rangle,$$

ello exige que siga siendo válida la misma relación de superposición entre los estados de movimiento durante todo el tiempo en que el sistema permanezca inalterado. Por tanto, debe conducirnos a la misma ecuación entre los kets correspondientes a dichos estados en un instante t (dentro del intervalo de tiempo en que no hay alteración), o sea, a la ecuación

$$|Rt\rangle = c_1|At\rangle + c_2|Bt\rangle,$$

siempre que elijamos de forma conveniente los factores numéricos arbitrarios que pueden multiplicar a estos kets. De aquí se sigue que los $|Pt\rangle$ son funciones lineales de los $|Pt_0\rangle$ y que cada $|Pt\rangle$ es el resultado de aplicar un cierto operador lineal a $|Pt_0\rangle$. O sea

$$\boxed{|Pt\rangle = T|Pt_0\rangle,} \tag{1}$$

donde T es un operador lineal independiente de P y dependiente sólo de t (y t_0).

Supongamos ahora que cada $|Pt\rangle$ tenga la misma longitud que el correspondiente $|Pt_0\rangle$. No siempre es posible escoger los factores numéricos arbitrarios por los cuales, según hemos visto, deben ir multiplicados los $|Pt\rangle$ sin destruir la dependencia lineal de los $|Pt\rangle$ con los $|Pt_0\rangle$; de esta forma, la nueva hipótesis es de naturaleza física, y no una simple cuestión de notación. Constituye una generalización del principio de superposición. La arbitrariedad en $|Pt\rangle$ se convierte así en un simple factor de fase que, si se quiere conservar la dependencia lineal de los $|Pt\rangle$ con los $|Pt_0\rangle$, ha de ser independiente de P. De la condición de que la longitud $c_1|Pt\rangle + c_2|Qt\rangle$ sea igual a la de $c_1|Pt_0\rangle + c_2|Qt_0\rangle$ para cualesquiera números complejos c_1 y c_2, podemos deducir que

$$\langle Qt|Pt\rangle = \langle Qt_0|Pt_0\rangle. \tag{2}$$

La conexión entre los $|Pt\rangle$ y los $|Pt_0\rangle$ es análoga formalmente a la que ligaba los kets trasladados y los no trasladados en § 25, si bien ahora se trata de un proceso de traslación en el tiempo en vez de en el espacio. Las ecuaciones (1) y (2) juegan el papel de las ecuaciones (59) y (60) de § 25. Podemos desarrollar las consecuencias de estas ecuaciones de la misma manera que hicimos en § 25 y deducir que T contiene un factor numérico arbitrario de módulo unidad y verifica

$$TT = 1, \tag{3}$$

en correspondencia con (62) de § 25; luego, T *es unitario*. Pasamos al caso

infinitesimal haciendo $t \to t_0$ y suponiendo por continuidad física que existe el límite

$$\lim_{t \to t_0} \frac{|Pt\rangle - |Pt_0\rangle}{t - t_0}$$

Dicho límite es la derivada de $|Pt_0\rangle$ respecto a t_0. Según (1), esto equivale a

$$\frac{d|Pt_0\rangle}{dt_0} = \left\{ \lim_{t \to t_0} \frac{T - 1}{t - t_0} \right\} |Pt_0\rangle. \tag{4}$$

Igual que en (64) de § 25, el operador límite que aparece aquí es un operador lineal imaginario puro y está indeterminado en un número arbitrario aditivo imaginario puro. Llamando H a este operador límite multiplicado por $i\hbar$, o mejor $H(t_0)$ ya que puede depender de t_0, la ecuación (4) se transforma, cuando la escribimos para un t general, en

$$i\hbar \frac{d|Pt\rangle}{dt} = H(t)|Pt\rangle. \tag{5}$$

La ecuación (5) expresa la ley general de variación con el tiempo del ket correspondiente al estado en un instante cualquiera. Es *la ecuación del movimiento en imagen de Schrödinger*. En ella aparece un operador lineal real $H(t)$, que ha de ser característico del sistema dinámico en cuestión. Supondremos que *$H(t)$ es la energía total del sistema*. Se pueden dar dos justificaciones de esta hipótesis, (i) la analogía que existe con la mecánica clásica que será desarrollada en la sección próxima, y (ii) $H(t)$ es igual a un operador de traslación en el tiempo multiplicado por $i\hbar$, similar a los operadores de traslación en las direcciones x, y, z de § 25, que en correspondencia con (69) de § 25 induce a tomar $H(t)$ igual a la energía total, pues la teoría de la relatividad sitúa la energía en la misma relación respecto al tiempo que el momento respecto a la distancia.

Por razones físicas, supondremos que la energía total de un sistema es siempre un observable. Para un sistema aislado es una constante y se puede escribir H. Muchas veces, aunque no sea una constante, la escribiremos simplemente H, sobreentendiendo su dependencia temporal. Si la energía depende de t, significa que actúan fuerzas externas sobre el sistema. Hay que distinguir entre una acción de este tipo y una alteración causada por un proceso de observación, pues la primera es compatible con la causalidad y las ecuaciones del movimiento y la segunda no.

Podemos establecer una relación entre $H(t)$ y el T de la ecuación (1). Así

$$i\hbar \frac{dT}{dt} |Pt_0\rangle = H(t)T|Pt_0\rangle.$$

Como $|Pt_0\rangle$ puede ser un ket cualquiera, tenemos

$$i\hbar \frac{dT}{dt} = H(t)T. \tag{6}$$

La ecuación (5) tiene mucha importancia en problemas prácticos, donde, a menudo, se utiliza unida a una representación. Introduciendo una representación mediante un conjunto completo de observables diagonales ξ que conmutan, y haciendo $\langle\xi'|Pt\rangle$ igual a $\psi(\xi't)$, tenemos, al pasar a la notación del ket standard,

$$|Pt\rangle = \psi(\xi t)\rangle.$$

La ecuación (5) se transforma en

$$i\hbar \frac{\partial}{\partial t}\psi(\xi t)\rangle = H\psi(\xi t)\rangle. \tag{7}$$

La ecuación (7) se conoce con el nombre de *ecuación de onda de Schrödinger*, y sus soluciones $\psi(\xi t)$ son *funciones de onda dependientes del tiempo*. Cada solución corresponde a un estado de movimiento del sistema, y el cuadrado de su módulo da la probabilidad de que los ξ tengan valores determinados en un instante t dado. Para los sistemas que se pueden describir mediante un conjunto de coordenadas canónicas podemos elegir una representación de Schrödinger y tomar entonces H como un operador de derivación de acuerdo con (42) de § 22.

28. *Ecuaciones del movimiento en imagen de Heisenberg*

En la sección precedente hemos establecido una imagen de los estados de movimiento no alterados, en la que cada uno de ellos correspondía a un ket en movimiento, y el estado en un instante cualquiera correspondía el ket en dicho instante. Esta descripción se designa *imagen de Schrödinger*. Aplicaremos a estos kets la transformación unitaria que hace pasar cada ket $|a\rangle$ al

$$|a^*\rangle = T^{-1}|a\rangle. \tag{8}$$

Esta transformación es de la forma dada por (75) de § 26, con T^{-1} para U, pero *depende del tiempo t*, ya que T depende de t. La imagen de esta transformación es la aplicación de un movimiento continuo de todo el espacio de los kets (que consiste en rotaciones y deformaciones uniformes). Un ket fijo inicialmente pasa a estar en movimiento, movimiento que viene dado por (8) con $|a\rangle$ independiente de t. En cambio, un ket que inicialmente estaba en movimiento siguiendo a un estado de movimiento no alterado, o sea, moviéndose de acuerdo con la ecuación (1), pasa a estar fijo,

ya que al sustituir $|Pt\rangle$ por $|a\rangle$ en (8) obtenemos $|a^*\rangle$ independiente de t. De esta forma, *la transformación conduce al reposo los kets correspondientes a los estados de movimiento no alterados.*

Debemos aplicar también la transformación unitaria a los bras y a los operadores lineales, de forma que las ecuaciones entre las diversas cantidades permanezcan invariantes. La transformación aplicada a los bras viene dada por la imaginaria conjugada de (8), y la aplicada a los operadores lineales está dada por (70) de § 26 con T^{-1} para U, es decir,

$$\alpha^* = T^{-1}\alpha T. \tag{9}$$

Un operador lineal que inicialmente está fijo se transforma en general en un operador lineal en movimiento. Ahora bien, una variable dinámica corresponde a una operador lineal que está inicialmente fijo (pues no depende de t de ninguna forma), y por tanto, después de la transformación le corresponde un operador lineal en movimiento. Así, la transformación nos conduce a una nueva imagen del movimiento, en la cual los estados corresponden a vectores fijos y las variables dinámicas a operadores lineales en movimiento. Esta descripción constituye la *imagen de Heisenberg.*

La condición física del sistema dinámico en cualquier instante implica una relación entre las variables dinámicas y el estado; la variación de la condición física en el tiempo puede ser atribuida o bien a un cambio en el estado manteniendo fijas las variables dinámicas, que nos lleva a la imagen de Schrödinger, o bien a un cambio en las variables dinámicas en el que permanecen fijos los estados, que constituye la imagen de Heisenberg.

En la imagen de Heisenberg hay ecuaciones de movimiento para las variables dinámicas. Tomemos una variable dinámica correspondiente al operador lineal fijo v en la imagen de Schrödinger. En la imagen de Heisenberg corresponderá a un operador lineal en movimiento que designaremos v_t en vez de v^* para poner de manifiesto su dependencia de t, y que vendrá dado por

$$v_t = T^{-1}vT \tag{10}$$

o bien

$$Tv_t = vT.$$

Derivando respecto a t, tenemos

$$\frac{dT}{dt}v_t + T\frac{dv_t}{dt} = v\frac{dT}{dt}.$$

Con la ayuda de (6), resulta

$$HTv_t + i\hbar T\frac{dv_t}{dt} = vHT$$

o sea

$$i\hbar \frac{dv_t}{dt} = T^{-1}vHT - T^{-1}HTv_t$$
$$= v_tH_t - H_tv_t, \tag{11}$$

donde

$$H_t = T^{-1}HT. \tag{12}$$

La ecuación (11) se puede escribir en notación de P.B.

$$\frac{dv_t}{dt} = [v_t, H_t]. \tag{13}$$

Las ecuaciones (11) o (13) indican cómo varía con el tiempo cualquier variable dinámica en la imagen de Heisenberg, y constituyen *las ecuaciones del movimiento en imagen de Heisenberg*. Estas ecuaciones del movimiento están determinadas por el operador lineal H que aparece en las ecuaciones del movimiento en imagen de Schrödinger y corresponde a la energía en la imagen de Heisenberg. A las variables dinámicas en imagen de Heisenberg, en la que éstas varían con el tiempo, las denominaremos *variables dinámicas de Heisenberg*, para distinguirlas de las variables dinámicas fijas de la imagen de Schrödinger, que llamaremos *variables dinámicas de Schrödinger*. Cada variable dinámica de Heisenberg está relacionada con la correspondiente variable dinámica de Schrödinger a través de la ecuación (10). Como esta ecuación es una transformación unitaria, todas las relaciones algebraicas y funcionales son idénticas para las dos clases de variables dinámicas. Para $t = t_0$ es $T = 1$, de donde $v_{t_0} = v$, y cualquier variable dinámica de Heisenberg en el instante t_0 es igual a la correspondiente variable dinámica de Schrödinger.

La ecuación (13) puede compararse con la mecánica clásica, donde también tenemos variables dinámicas que varían con el tiempo. Las ecuaciones del movimiento de la mecánica clásica se pueden escribir en la forma de Hamilton

$$\frac{dq_r}{dt} = \frac{\partial H}{\partial p_r}, \qquad \frac{dp_r}{dt} = -\frac{\partial H}{dq_r}, \tag{14}$$

donde las q_r y las p_r son un conjunto de coordenadas canónicas y H es la energía expresada en función de ellas y posiblemente también de t. La energía expresada en esta forma recibe el nombre de *hamiltoniano*. Las ecuaciones (14) dan, para cualquier función v de las q_r y las p_r que no contenga explícitamente el tiempo t,

$$\frac{dv}{dt} = \sum_r \left\{ \frac{\partial v}{\partial q_r}\frac{dq_r}{dt} + \frac{\partial v}{\partial p_r}\frac{dp_r}{dt} \right\}$$
$$= \sum_r \left\{ \frac{\partial v}{\partial q_r}\frac{\partial H}{\partial p_r} - \frac{\partial v}{\partial p_r}\frac{\partial H}{\partial q_r} \right\}$$
$$= [v, H], \tag{15}$$

según la definición clásica de P.B. dada por la ecuación (1) de § 21. Esta ecuación es de la misma forma que la ecuación (13) en la teoría cuántica. Así hemos obtenido una analogía entre las ecuaciones del movimiento clásicas en forma de Hamilton y las ecuaciones del movimiento cuánticas en imagen de Heisenberg. Esta analogía nos da una justificación de la hipótesis de que el operador lineal H introducido en la sección precedente es la energía del sistema en mecánica cuántica.

En mecánica clásica, un sistema dinámico está definido matemáticamente cuando conocemos el hamiltoniano, es decir, cuando conocemos la energía expresada como función de un conjunto de coordenadas canónicas, pues ello es suficiente para determinar las ecuaciones del movimiento. En mecánica cuántica, un sistema dinámico está definido matemáticamente cuando conocemos la energía expresada en función de variables dinámicas cuyas relaciones de conmutación son conocidas, pues ello basta para determinar las ecuaciones del movimiento tanto en imagen de Schrödinger como en imagen de Heisenberg. Necesitamos tener, o bien H expresado en función de las variables de Schrödinger, bien H_t expresado en función de las correspondientes variables dinámicas de Heisenberg, siendo la dependencia funcional la misma en ambos casos. A la energía expresada en estos términos la denominaremos *hamiltoniano* del sistema dinámico en mecánica cuántica, manteniendo así la analogía con la teoría clásica.

Todo sistema de la mecánica cuántica tiene siempre un hamiltoniano, tanto si es un sistema que tiene análogo clásico y se puede describir mediante un conjunto de coordenadas canónicas como si no lo es. Pero si el sistema tiene un análogo clásico, su conexión con la mecánica clásica es particularmente estrecha, y por lo general se puede suponer que el hamiltoniano es la misma función de las coordenadas y momentos canónicos en la teoría cuántica que en la teoría clásica.† Desde luego esto entrañaría una dificultad si el hamiltoniano clásico comprendiera un producto de factores cuyos análogos cuánticos no conmutaran, ya que no se sabría en qué orden poner estos factores en el hamiltoniano cuántico, sin embargo, esto no ocurre para la mayoría de sistemas dinámicos elementales cuyo estudio tiene importancia en física atómica. En consecuencia, podemos utilizar ampliamente el mismo lenguaje para describir sistemas dinámicos en la teoría cuántica que en la teoría clásica (por ejemplo, refererirnos a partículas con masas dadas moviéndose en campos de fuerza dados), y dado un sistema en mecánica clásica, en general, podemos dar una significación de 'el mismo' sistema en mecánica cuántica.

La ecuación (13) se verifica para cualquier función v_t de las variables dinámicas de Heisenberg que no sea función explícita del tiempo, es decir,

† Esta hipótesis ha tenido éxito en la práctica únicamente cuando se ha aplicado a coordenadas y momentos dinámicos referidos a un sistema de ejes cartesianos y no a coordenadas curvilíneas más generales.

para cualquier operador lineal v que en la imagen de Schrödinger sea constante. Esto demuestra que una tal función v_t es constante si conmuta con H_t, o lo que es lo mismo, si v conmuta con H. En ese caso tenemos

$$v_t = v_{t_0} = v,$$

y decimos que v_t o v es una *constante del movimiento*. Es necesario que v conmute con H en cualquier instante, que en general sólo es posible si H es constante. En este caso podemos sustituir H por v en (13) y deducir que H_t es constante, lo cual demuestra que el mismo H es una constante del movimiento. Así pues, si el hamiltoniano es constante en la imagen de Schrödinger, también lo es en la imagen de Heisenberg.

Para un sistema aislado sobre el que no actúan fuerzas exteriores, siempre existen algunas constantes del movimiento. Una de ellas es la energía total o hamiltoniano. Otras, se derivan de la teoría de traslaciones de § 25. Es evidente desde el punto de vista físico, que la energía total debe permanecer invariante si todas las variables dinámicas se trasladan de un mismo modo, y por tanto, la ecuación (63) de § 25 debe verificarse para $v_d = v = H$. Así pues, D conmuta con H y es una constante del movimiento. Pasando al caso de una traslación infinitesimal, vemos que los operadores de traslación d_x, d_y, d_z son constantes del movimiento, y por lo tanto, según (69) de § 25, el momento total es una constante del movimiento. Asimismo, la energía total permanecerá invariante si llevamos a cabo un giro de todas las variables dinámicas. De ello deduciremos en § 35 que el momento angular total es una constante del movimiento. *Las leyes de conservación de energía, momento y momento angular son válidas también para un sistema aislado en la imagen de Heisenberg en mecánica cuántica, igual que ocurría en mecánica clásica.*

Hemos dado dos formas de las ecuaciones del movimiento en mecánica cuántica. De ellas, la de Schrödinger es la más empleada en problemas prácticos, pues da lugar a las ecuaciones más sencillas. Las incógnitas en la ecuación de onda de Schrödinger son los números que constituyen el representante de un ket, mientras que la ecuación del movimiento de Heisenberg para una variable dinámica, si se proyecta sobre una representación, tiene como incógnitas los números que constituyen el representante de la variable dinámica. Los últimos son mucho más numerosos y, por lo tanto, más difíciles de calcular que las incógnitas de Schrödinger. El valor de las ecuaciones del movimiento en imagen de Heisenberg estriba en que nos proporcionan una analogía inmediata con la mecánica clásica, que nos permite ver cómo se pueden traducir a la teoría cuántica las diversas características de la teoría clásica, tales como las leyes de conservación a las que hemos aludido antes.

(*) el momento lineal se conserva
y " " angular " " " "

29. *Estados estacionarios*

Vamos a considerar aquí sistemas dinámicos cuya energía permanece constante. Para este caso son válidas algunas relaciones especialmente sencillas. Se puede integrar † la ecuación (6), y teniendo en cuenta la condición inicial de que $T = 1$ para $t = t_0$, da

$$T = e^{-iH(t-t_0)/\hbar},$$

Este resultado sustituido en (1) da

$$|Pt\rangle = e^{-iH(t-t_0)/\hbar}|Pt_0\rangle, \tag{16}$$

que es la integral de la ecuación del movimiento de Schrödinger, y sustituido en (10) da

$$v_t = e^{iH(t-t_0)\hbar} v e^{-iH(t-t_0)/\hbar}, \tag{17}$$

que es la integral de la ecuación del movimiento de Heisenberg (11), siendo H_t igual a H. Hemos obtenido así fácilmente soluciones de las ecuaciones del movimiento. Sin embargo, estas soluciones no tienen mucho valor práctico debido a la dificultad que lleva consigo el cálculo del operador $e^{-iH(t-t_0)/\hbar}$, a menos que H sea particularmente simple, y generalmente para aplicaciones prácticas es necesario volver a considerar la ecuación de onda de Schrödinger.

Consideremos un estado de movimiento que en el instante t_0 sea un autoestado de la energía. El ket $|Pt_0\rangle$ que le corresponde en dicho instante será un autoket de H. Si H' es el autovalor al que pertenece, la ecuación (16) da

$$|Pt\rangle = e^{-iH'(t-t_0)/\hbar}|Pt_0\rangle,$$

que indica que $|Pt\rangle$ difiere de $|Pt_0\rangle$ tan sólo en un factor de fase. Por tanto, el estado continúa siendo siempre un autoestado de la energía, y además no varía con el tiempo, puesto que la dirección de $|Pt\rangle$ no varía con el tiempo. Estos estados se denominan *estados estacionarios*. La probabilidad de obtener cualquier resultado particular en una observación de él es independiente del tiempo en que se lleva a cabo la observación. Teniendo en cuenta nuestra hipótesis de que la energía es un observable, vemos que hay suficientes estados estacionarios para que un estado arbitrario sea dependiente de ellos.

La función de onda $\psi(\xi t)$ dependiente del tiempo, que representa un estado estacionario de energía H', variará con el tiempo de acuerdo con la ley

$$\psi(\xi t) = \psi_0(\xi)e^{-iH't/\hbar}, \tag{18}$$

---

† La integración se puede efectuar como si H fuese una variable algebraica ordinaria en vez de un operador lineal, porque en la operación no hay ninguna cantidad que no conmute con H.

y la ecuación de onda de Schrödinger (7) se reduce para ella a

$$H'\psi_0\rangle = H\psi_0\rangle. \tag{19}$$

Esta ecuación afirma simplemente que el estado representado por ψ_0 es un autoestado de H. A una función ψ_0 que satisfaga (19) le llamaremos autofunción de H perteneciente al autovalor H'.

En la imagen de Heisenberg, los estados estacionarios corresponden a autovectores fijos de la energía. Podemos establecer una representación en la cual todos los vectores básicos sean autovectores de la energía y que, en consecuencia, correspondan a estados estacionarios en la imagen de Heisenberg. A esta representación la denominaremos *representación de Heisenberg*. El primer modelo de mecánica cuántica ideado por Heisenberg en 1925 estaba expresado en una representación de este tipo. La energía es diagonal en dicha representación. Cualquier otra variable dinámica diagonal debe conmutar con la energía, y por tanto, es una constante del movimiento. El problema de establecer una representación de Heisenberg se reduce entonces al problema de encontrar un conjunto completo de observables que conmutan, cada uno de los cuales es una constante del movimiento, y después diagonalizar estos observables. La energía debe ser una función de estos observables según el teorema 2 de § 19. Algunas veces es conveniente tomar la misma energía como uno de ellos.

Designemos con α el conjunto completo de observables que conmutan en una representación de Heisenberg, con lo que los vectores básicos se escribirán $\langle\alpha'|$, $|\alpha''\rangle$. La energía es una función de estos observables α, o sea $H = H(\alpha)$. De (17) obtenemos

$$\begin{aligned}\langle\alpha'|v_t|\alpha''\rangle &= \langle\alpha'|e^{iH(t-t_0)/\hbar}ve^{-iH(t-t_0)/\hbar}|\alpha''\rangle \\ &= e^{i(H'-H'')(t-t_0)/\hbar}\langle\alpha'|v|\alpha''\rangle,\end{aligned} \tag{20}$$

donde $H' = H(\alpha')$ y $H'' = H(\alpha'')$. El factor $\langle\alpha'|v|\alpha''\rangle$ que aparece en el segundo miembro es independiente de t, y es un elemento de la matriz que representa al operador lineal fijo v. La fórmula (20) muestra cómo varían con el tiempo los elementos de matriz de Heisenberg de cualquier variable dinámica de Heisenberg, y hace que v_t satisfaga la ecuación del movimiento (11), como puede comprobarse fácilmente. La variación dada por (20) es simplemente periódica con una frecuencia

$$|H' - H''|/2\pi\hbar = |H' - H''|/h, \tag{21}$$

que sólo depende de la diferencia de energías entre los dos estados a los que se refiere el elemento de matriz. Este resultado está íntimamente ligado a la ley de combinación de la espectroscopia y a la condición de Bohr de las frecuencias, según la cual (21) es la frecuencia de la radiación electromagnética emitida o absorbida cuando el sistema efectúa una transición

entre dos estados estacionarios α' y α'', bajo la influencia de una radiación, siendo los autovalores de H los niveles de energía de Bohr. Estas cuestiones serán tratadas en § 45.

30. *La partícula libre*

La aplicación fundamental más sencilla de la mecánica cuántica consiste en considerar un sistema formado simplemente por una partícula libre, o sea, una partícula no sometida a ninguna fuerza. Utilizaremos como variables dinámicas las tres coordenadas cartesianas, x, y, z y sus momentos conjugados p_x, p_y, p_z. El hamiltoniano es igual a la energía cinética de la partícula, es decir,

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) \tag{22}$$

de acuerdo con la mecánica de Newton, siendo m la masa. Esta fórmula es válida únicamente si la velocidad de la partícula es pequeña comparada con c, velocidad de la luz. Para una partícula que se mueve con gran velocidad, caso que se presenta a menudo en la teoría atómica, (22) debe ser reemplazada por la fórmula relativista

$$H = c(m^2c^2 + p_x^2 + p_y^2 + p_z^2)^{\frac{1}{2}}. \tag{23}$$

Para pequeños valores de p_x, p_y y p_z, (23) se convierte en (22) salvo el término constante mc^2 que corresponde a la energía en reposo de la partícula en la teoría de la relatividad, y que no influye en la ecuación del movimiento. Las fórmulas (22) y (23) se pueden traducir directamente a la teoría cuántica, sobreentendiéndose que la raíz cuadrada que aparece en (23) es la raíz cuadrada positiva definida al final de § 11. El término constante mc^2 en que difiere (23) de (22) para pequeños valores de p_x, p_y, p_z no puede tener efecto físico alguno, pues el hamiltoniano en la teoría cuántica, tal como lo hemos introducido en § 27, está indeterminado en una constante real aditiva.

Aquí emplearemos la fórmula más exacta (23). Resolveremos en primer lugar las ecuaciones del movimiento de Heisenberg. Según las condiciones cuánticas (9) de § 21, p_x conmuta con p_y y p_z, y por tanto, según el teorema 1 de § 19 extendido a un conjunto de observables que conmutan, p_x conmuta con cualquier función de p_x, p_y y p_z, y por tanto, con H. Luego p_x es una constante del movimiento. Análogamente p_y y p_z son también constantes del movimiento. Estos resultados son los mismos que en la teoría clásica. Por otra parte, la ecuación del movimiento para una coordenada, por ejemplo x_t, según (11), es

$$i\hbar\,\dot{x}_t = i\hbar\frac{dx_t}{dt} = x_t\,c(m^2c^2+p_x^2+p_y^2+p_z^2)^{\frac{1}{2}} - c(m^2c^2+p_x^2+p_y^2+p_z^2)^{\frac{1}{2}}x_t.$$

Podemos calcular el segundo miembro mediante la fórmula (31) de § 22 cambiando los papeles respectivos de las coordenadas y de los momentos, de forma que se tenga

$$q_r f - f q_r = i\hbar\,\partial f/\partial p_r, \tag{24}$$

siendo f una función de las p_r. Esto da

$$\left.\begin{aligned} \dot{x}_t &= \frac{\partial}{\partial p_x}c(m^2c^2+p_x^2+p_y^2+p_z^2)^{\frac{1}{2}} = \frac{c^2p_x}{H}. \\ \text{Análogamente}\qquad \dot{y}_t &= \frac{c^2p_y}{H}, \qquad \dot{z}_t = \frac{c^2p_z}{H}. \end{aligned}\right\} \tag{25}$$

La magnitud de la velocidad es

$$v = (\dot{x}_t^2+\dot{y}_t^2+\dot{z}_t^2)^{\frac{1}{2}} = c^2(p_x^2+p_y^2+p_z^2)^{\frac{1}{2}}/H. \tag{26}$$

Las ecuaciones (25) y (26) son exactamente las mismas que en la teoría clásica.

Consideremos un estado que sea autoestado de los momentos perteneciente a los autovalores p'_x, p'_y, p'_z. Este estado tiene que ser un autoestado del hamiltoniano perteneciente al autovalor

$$H' = c(m^2c^2+p_x'^2+p_y'^2+p_z'^2)^{\frac{1}{2}}, \tag{27}$$

y será por tanto un estado estacionario. Los valores posibles de H' son todos los números desde mc^2 a ∞, como en la teoría clásica. La función de onda $\psi(xyz)$, que representa este estado en un instante cualquiera en la representación de Schrödinger, debe satisfacer

$$p'_x\psi(xyz)\rangle = p_x\psi(xyz)\rangle = -i\hbar\frac{\partial\psi(xyz)}{\partial x}\rangle,$$

con ecuaciones análogas para p_y y p_z. Estas ecuaciones indican que $\psi(xyz)$ es de la forma

$$\psi(xyz) = a e^{i(p'_x x+p'_y y+p'_z z)/\hbar}, \tag{28}$$

donde q es independiente de x, y, z. De (18) resulta que la función de onda $\Psi(xyzt)$ dependiente del tiempo es de la forma

$$\psi(xyzt) = a_0\, e^{i(p'_x x+p'_y y+p'_z z-H't)/\hbar}, \tag{29}$$

donde a_0 es independiente de x, y, z y t.

La función (29) de x, y, z y t representa ondas planas en el espacio-

tiempo. Vemos, a la luz de este ejemplo, la conveniencia de los términos 'función de onda' y 'ecuación de onda. La frecuencia de las ondas es

$$\nu = H'/h, \tag{30}$$

y su longitud de onda es

$$\lambda = h/(p'^2_x + p'^2_y + p'^2_z)^{\frac{1}{2}} = h/P', \tag{31}$$

donde P' es la longitud del vector (p'_x, p'_y, p'_z), y el movimiento tiene lugar en la dirección del vector (p'_x, p'_y, p'_z) con la velocidad

$$\lambda\nu = H'/P' = c^2/v', \tag{32}$$

donde v' es la velocidad de la partícula correspondiente al momento (p'_x, p'_y, p'_z) dada por la fórmula (26). Es fácil ver que las ecuaciones (30), (31) y (32) se verifican en todos los sistemas de referencia de Lorentz, pues de hecho, la expresión del segundo miembro de (29) es un invariante relativista, siendo p'_x, p'_y, p'_z y H' las componentes de un cuadrivector. Estas propiedades de invariancia relativista condujeron a de Broglie, antes del descubrimiento de la mecánica cuántica, a postular la existencia de ondas de la forma (29) asociadas al movimiento de cualquier partícula. Se conocen por tanto como *ondas de de Broglie.*

En el caso límite cuando se hace tender la masa m a cero, la velocidad clásica v de la partícula se hace igual a c, y por tanto, según (32), la velocidad de la onda también se hace c. Las ondas son semejantes entonces a las ondas luminosas asociadas a un fotón, con la diferencia de que no contienen información alguna sobre la polarización, y que llevan consigo un exponencial complejo en vez de senos y cosenos. Las fórmulas (30) y (31) siguen siendo válidas y relacionan la frecuencia de las ondas luminosas con la energía del fotón y la longitud de onda de las ondas luminosas con el momento del fotón.

Para el estado representando por (29), la probabilidad de encontrar la partícula en cualquier pequeño volumen cuando se lleva a cabo una observación de su posición es independiente del lugar en que se encuentra el volumen. Esto proporciona un ejemplo del principio de incertidumbre de Heisenberg, ya que se trata de un estado para el cual el momento está dado con toda exactitud y para el cual, en consecuencia, la posición está completamente indeterminada. Un tal estado es, desde luego, un caso límite que no ocurre nunca en la práctica. Los estados que comúnmente encontramos en la práctica son los representados por paquetes de ondas, que se pueden construir por la superposición de un conjunto de ondas del tipo (29) pertenecientes a valores ligeramente diferentes de (p'_x, p'_y, p'_z), como vimos en § 24. La fórmula ordinaria en hidrodinámica de la velocidad de un tal paquete de ondas, es decir, la *velocidad de grupo* de las ondas, es

$$\frac{d\nu}{d(1/\lambda)} \tag{33}$$

que, según (30) y (31), da

$$\frac{dH'}{dP'} = c\frac{d}{dP'}(m^2c^2 + P'^2)^{\frac{1}{2}} = \frac{c^2P'}{H'} = v'. \tag{34}$$

Ésta es precisamente la velocidad de la partícula. El paquete de ondas se mueve en la misma dirección y con la misma velocidad que la partícula en mecánica clásica.

31. *Movimiento de paquetes de ondas*

El último resultado que hemos deducido para una partícula libre constituye un ejemplo de un principio más general. Para todo sistema con análogo clásico, los estados para los que la mecánica clásica constituye una aproximación válida están representados en mecánica cuántica por paquetes de ondas, para los cuales todas las coordenadas y momentos tienen aproximadamente valores numéricos determinados, cuyo grado de determinación viene limitado por el principio de incertidumbre de Heisenberg. Pero la ecuación de Schrödinger determina cómo varían con el tiempo dichos paquetes de ondas, y para que la mecánica clásica siga siendo una descripción válida es necesario que el paquete de ondas conserve la forma de paquete de ondas y que se mueva según las leyes de la dinámica clásica. Vamos a comprobar que es así.

Sea un sistema dinámico que tenga un análogo clásico, y sea $H(q_r, p_r)$ $(r = 1, 2, \ldots, n)$ su hamiltoniano. El sistema dinámico clásico correspondiente tendrá el hamiltoniano $H_c(q_r, p_r)$ que se obtiene sustituyendo las q_r y p_r de $H(q_r, p_r)$ por variables algebraicas ordinarias y haciendo que $\hbar \to 0$ si es que interviene en $H(q_r, p_r)$. Es evidente que el hamiltoniano clásico H_c ha de ser una función real de sus variables. Generalmente es una función cuadrática respecto a los momentos p_r pero no siempre (la teoría relativista de una partícula libre es un ejemplo de ello). La demostración que vamos a dar a continuación es válida para cualquier función algebraica H_c de las p_r.

Supondremos que la función de onda dependiente del tiempo en la representación de Schrödinger es de la forma

$$\psi(qt) = Ae^{iS/\hbar}, \tag{35}$$

donde A y S son dos funciones reales de las q_r y de t que no varían muy rápidamente. Así pues, la función de onda tiene forma de ondas cuya amplitud y fase son respectivamente A y S. La ecuación de onda de Schrödinger (7) da

$$i\hbar\frac{\partial}{\partial t}Ae^{iS/\hbar}\rangle = H(q_r, p_r)Ae^{iS/\hbar}\rangle$$

o bien

$$\left\{ i\hbar \frac{\partial A}{\partial t} - A \frac{\partial S}{\partial t} \right\} \rangle = e^{-iS/\hbar} H(q_r, p_r) A e^{iS/\hbar} \rangle. \tag{36}$$

Evidentemente $e^{-iS/\hbar}\rangle$ es un operador lineal unitario que puesto en lugar de U en la ecuación (70) de § 26 nos determina una transformación unitaria. Las q_r permanecen invariantes en dicha transformación; cada p_r se transforma con ayuda de (31) de § 22 en

$$e^{-iS/\hbar} p_r e^{iS/\hbar} = p_r + \partial S/\partial q_r,$$

y H se transforma en

$$e^{-iS/\hbar} H(q_r, p_r) e^{iS/\hbar} = H(q_r, p_r + \partial S/\partial q_r),$$

puesto que las relaciones algebraicas se conservan en la transformación. Por tanto (36) se convierte en

$$\left\{ i\hbar \frac{\partial A}{\partial t} - A \frac{\partial S}{\partial t} \right\} \rangle = H\left(q_r, p_r + \frac{\partial S}{\partial q_r}\right) A \rangle. \tag{37}$$

Supongamos ahora que $\hbar$ puede considerarse pequeña, y despreciemos los términos en que figure $\hbar$ en (37). Debemos despreciar entonces en (37) las p_r que figuran en H, pues cada p_r equivale al operador $-i\hbar\partial/\partial q_r$ aplicado a las funciones de las q_r que figuran a la derecha. Los términos restantes nos dan

$$-\frac{\partial S}{\partial t} = H_c\left(q_r, \frac{\partial S}{\partial q_r}\right). \tag{38}$$

Ésta es una ecuación diferencial que debe verificar la función de fase S. La ecuación está determinada por la función de Hamilton clásica H_c, y en la dinámica clásica se conoce con el nombre de *ecuación de Hamilton-Jacobi.* Admite soluciones reales para S, y por tanto, la hipótesis de la forma ondulatoria de (35) no nos lleva a ninguna contradicción.

Para obtener una ecuación en que figure A, hemos de conservar en (37) los términos lineales en $\hbar$ y ver qué se obtiene con ellos. El cálculo directo de dichos términos en el caso de una función H general es bastante difícil, y para obtener el resultado que buscamos resulta más fácil multiplicar ambos miembros de (37) por el bra $\langle Af$, donde f es una función real de las q_r arbitraria. Así resulta

$$\langle Af \left\{ i\hbar \frac{\partial A}{\partial t} - A \frac{\partial S}{\partial t} \right\} \rangle = \langle Af H\left(q_r, p_r + \frac{\partial S}{\partial q_r}\right) A \rangle.$$

La ecuación compleja conjugada de ésta es

$$\langle Af\left\{-i\hbar\frac{\partial A}{\partial t}-A\frac{\partial S}{\partial t}\right\}\rangle = \langle AH\left(q_r, p_r+\frac{\partial S}{\partial q_r}\right)fA\rangle.$$

Restándolas y dividiendo por $i\hbar$ obtenemos

$$2\langle Af\frac{\partial A}{\partial t}\rangle = \langle A\left[f,\ H\left(q_r, p_r+\frac{\partial S}{\partial q_r}\right)\right]A\rangle. \tag{39}$$

Tenemos que calcular el P.B.

$$[f,\ H(q_r, p_r+\partial S/\partial q_r)].$$

La hipótesis de que $\hbar$ puede ser considerada pequeña nos permite desarrollar $H(q_r,\ p_r+\partial S/\partial q_1)$ en serie de potencias de las p_r. Los términos de orden cero no contribuyen al valor del P.B. La forma más cómoda de calcular la contribución de los términos de primer orden en las p_r al valor del P.B. es utilizar la fórmula clásica (1) del § 21 (dicha fórmula es válida también en la teoría cuántica cuando u es independiente de las p_r y v es lineal en las p_r). La contribución total es

$$\sum_s \frac{\partial f}{\partial q_s}\left[\frac{\partial H(q_r, p_r)}{\partial p_s}\right]_{p_r=\partial S/\partial q_r}$$

donde el corchete con subíndices significa que hemos de sustituir, en la función de las q_r y las p_r, que figura dentro, cada p_r por $\partial S/\partial q_r$, con lo cual obtenemos una función únicamente de las q_r. Los términos de órdenes superiores en las p_r dan una contribución al valor del P.B. que para $\hbar\to 0$ se anula. Luego, despreciando los términos en que figura $\hbar$ en (39), que equivale a despreciar $\hbar^2$ en (37), (39) se convierte en

$$\langle f\frac{\partial A^2}{\partial t}\rangle = \langle A^2\sum_s \frac{\partial f}{\partial q_s}\left[\frac{\partial H_c(q_r, p_r)}{\partial p_s}\right]_{p_r=\partial S/\partial q_r}\rangle. \tag{40}$$

Por otro lado la fórmula (64) de § 20 nos da para $a(q)$ y $b(q)$ dos funciones de las q_r cualesquiera

$$\langle a(q)b(q)\rangle = \int a(q')dq'\,b(q'),$$

de donde

$$\langle a(q)\frac{\partial b(q)}{\partial q_r}\rangle = -\langle\frac{\partial a(q)}{\partial q_r}b(q)\rangle, \tag{41}$$

siempre que $a(q)$ y $b(q)$ verifiquen unas condiciones de contorno adecuadas tales como las que discutimos en §§ 22 y 23. Luego, (40) se puede escribir

$$\langle f \frac{\partial A^2}{\partial t} \rangle = -\langle f \sum_s \frac{\partial}{\partial q_s} \left\{ A^2 \left[\frac{\partial H_c(q_r, p_r)}{\partial p_s} \right]_{p_r = \partial S/\partial q_r} \right\} \rangle.$$

Puesto que esto es válido para toda función real f, se tiene

$$\frac{\partial A^2}{\partial t} = -\sum_s \frac{\partial}{\partial q_s} \left\{ A^2 \left[\frac{\partial H_c(q_r, p_r)}{\partial p_s} \right]_{p_r = \partial S/\partial q_r} \right\}. \tag{42}$$

Ésta es la ecuación que debe satisfacer la amplitud A de la función de onda. Para comprender su significado, consideremos un fluido que se mueve en el espacio de las variables q_r cuya densidad en cualquier punto de espacio sea igual a A^2 y cuya velocidad sea

$$\frac{dq_s}{dt} = \left[\frac{\partial H_c(q_r, p_r)}{\partial p_s} \right]_{p_r = \partial S/\partial q_r}. \tag{43}$$

En esta imagen, la ecuación (42) es precisamente la ecuación de conservación del fluido. El movimiento del fluido está determinado por la función S que verifica (38), y para cada solución de (38) existirá un posible movimiento.

Para una S dada consideremos una solución de (42) para la cual, en un instante inicial dado, la densidad A^2 sea nula en todo el espacio exterior a una pequeña región. Podemos suponer que dicha región se mueve siguiendo al fluido, es decir, con una velocidad dada por (43), y entonces la ecuación de conservación (42) exigirá que la densidad fuera de dicha región sea siempre nula. Existe un límite inferior del tamaño de dicha región, que viene impuesto por la aproximación que hemos hecho al despreciar $\hbar$ en (30). Dicha aproximación es válida únicamente si

$$\hbar \frac{\partial}{\partial q_r} A \ll \frac{\partial S}{\partial q_r} A,$$

o sea si

$$\frac{1}{A} \frac{\partial A}{\partial q_r} \ll \frac{1}{\hbar} \frac{\partial S}{\partial q_r}$$

que exige que al variar las q_r, A no varíe en una fracción apreciable de sí misma en un intervalo en el que S no haya variado en muchas veces $\hbar$, es decir, en un dominio de anchura superior a muchas longitudes de onda de la función de onda (35). Si se verifica dicha condición, la solución considerada es un paquete de ondas del tipo estudiado en § 24, y sigue siéndolo durante todo el tiempo.

Hemos obtenido así una función de onda que representa un estado de movimiento para el cual las coordenadas y los momentos tienen aproximadamente valores numéricos determinados durante todo el tiempo. Tales

estados de movimiento de la mecánica cuántica corresponden a los estados que considera la teoría clásica. El movimiento del paquete de ondas está determinado por las ecuaciones (38) y (43). De ellas, definiendo p_s igual a $\partial S/\partial q_s$, obtenemos

$$\begin{aligned}\frac{dp_s}{dt} = \frac{d}{dt}\frac{\partial S}{\partial q_s} &= \frac{\partial^2 S}{\partial t \partial q_s} + \sum_u \frac{\partial^2 S}{\partial q_u \partial q_s}\frac{\partial q_u}{dt} \\ &= -\frac{\partial}{\partial q_s} H_c\left(q_r, \frac{\partial S}{\partial q_r}\right) + \sum_u \frac{\partial^2 S}{\partial q_u \partial q_s}\frac{\partial H_c(q_r, p_r)}{\partial p_u} \\ &= -\frac{\partial H_c(q_r, p_r)}{\partial q_s} \qquad (44)\end{aligned}$$

en donde en la última línea las q_r y las p_r las hemos considerado independientes antes de la derivación parcial. Las ecuaciones (43) y (44) coinciden con las ecuaciones de movimientos clásicas puestas en forma de Hamilton e indican que el paquete de ondas se mueve según las leyes de la mecánica clásica. Vemos así cómo pueden deducirse las ecuaciones del movimiento clásicas como caso límite a partir de la teoría cuántica.

Mediante una solución más exacta de la ecuación de onda puede demostrarse que el grado de precisión con que pueden atribuirse simultáneamente valores numéricos a las coordenadas y momentos no puede permanecer constantemente en el límite permitido por el principio de Heisenberg — ecuación (36) de § 24 — y que si inicialmente ocurre así, en el transcurso del tiempo irá disminuyendo y el paquete de ondas se irá deformando.†

32. *El principio de acción* ‡

La ecuación (10) nos dice que las variables dinámicas de Heisenberg v_t en el instante t, están ligadas con sus valores v_{t_0} o v en el instante t_0, a través de una transformación unitaria. Las variables dinámicas de Heisenberg en el instante $t + \delta t$ están ligadas con sus valores en el instante t a través de una transformación unitaria infinitesimal como muestra la ecuación de movimiento (11) o (13), que relaciona $v_{t+\delta t}$ y v_t en la forma (79) u (80) de § 26 con H_t en lugar de F y $\delta t/\hbar$ en vez de ε. Por tanto, la variación con el tiempo de las variables dinámicas de Heisenberg puede ser considerada como el desarrollo continuo de una transformación unitaria.

† Véase KENNARD, *Z. f. Physik*, 44 (1927), 344; DARWIN, *Proc. Roy. Soc. A*, 117 (1927), 258.

‡ Esta sección puede ser omitida por el estudiante que no esté especialmente interesado en dinámica superior.

En mecánica clásica las variables dinámicas en el instante $t + \delta t$ están relacionadas con los valores en el instante t mediante una transformación de contacto infinitesimal, y el movimiento global puede ser considerado como el desarrollo continuo de una transformación de contacto. Éste es el fundamento matemático de la analogía entre las ecuaciones del movimiento clásicas y cuánticas, y podemos desarrollarlo a fin de descubrir el análogo cuántico de todos los principales rasgos característicos de la teoría clásica de la dinámica.

Sea una representación en la que el conjunto completo de observables que conmutan ξ sea diagonal, siendo por tanto los bras básicos $\langle\xi'|$. Podemos introducir una segunda representación cuyos bras básicos sean

$$\langle\xi'^*| = \langle\xi'|T. \tag{45}$$

Los nuevos bras básicos son función del tiempo t y forman una representación móvil análoga a un sistema de referencia móvil en un espacio vectorial ordinario. Comparando (45) con la conjugada imaginaria de (8), vemos que los nuevos vectores básicos son precisamente los transformados en la imagen de Heisenberg de los vectores básicos originales de la imagen de Schrödinger y, en consecuencia, estarán ligados a las variables dinámicas de Heisenberg v_t del mismo modo que los vectores básicos de partida lo están a las variables dinámicas de Schrödinger v. En particular cada $\langle\xi'^*|$ ha de ser un autovector de los ξ_t perteneciente a los autovalores ξ'. Por tanto, podremos escribirlo $\langle\xi'_t|$, sobreentendiendo que los números ξ'_t son los mismos autovalores de los ξ_t que los ξ' son de los ξ. De (45) resulta

$$\langle\xi'_t|\xi''\rangle = \langle\xi'|T|\xi''\rangle, \tag{46}$$

que nos dice que la función de transformación coincide con el representante de T en la representación original.

Diferenciando (45) con respecto a t y empleando (6), con ayuda de (12), resulta

$$i\hbar\frac{d}{dt}\langle\xi'| = i\hbar\langle\xi'|\frac{dT}{dt} = \langle\xi'|HT = \langle\xi'|T$$

Multiplicando a la derecha por un ket $|a\rangle$ cualquiera que no sea función de t, resulta

$$i\hbar\frac{d}{dt}\langle\xi'_t|a\rangle = \langle\xi'_t|H_t|a\rangle = \int \langle\xi'_t|H_t|\xi''_t\rangle\, d\xi''_t\, \langle\xi''_t|a\rangle, \tag{47}$$

donde para concretar hemos considerado el caso en que los autovalores de las ξ son continuos. Por otro lado, la ecuación (5) escrita en forma de representantes es

$$i\hbar\frac{d}{dt}\langle\xi'|Pt\rangle = \int \langle\xi'|H|\xi''\rangle\, d\xi''\, \langle\xi''|Pt\rangle. \tag{48}$$

Como $\langle \xi'_t | H | \xi''_t \rangle$ es la misma función de las variables ξ'_t y ξ''_t que $\langle \xi' | H | \xi'' \rangle$ es de las ξ' y ξ'', las ecuaciones (47) y (48) son exactamente de la misma forma; las variables ξ'_t, ξ''_t de (47) juegan el papel de las variables ξ' y ξ'' de (48), y la función $\langle \xi'_t | a \rangle$ el de la función $\langle \xi' | Pt \rangle$. Así pues, podemos considerar (47) como una forma de la ecuación de onda de Schrödinger con la función de onda $\langle \xi'_t | a \rangle$ de las variables ξ'_t. De este modo *la ecuación de onda de Schrödinger aparece bajo otro aspecto, como la condición que han de verificar los representantes de los kets fijos que corresponden a los estados en la imagen de Heisenberg, en la representación móvil en la que las variables de Heisenberg ξ_t son diagonales.* La función $\langle \xi'_t | a \rangle$ debe su variación temporal al factor de la izquierda $\langle \xi'_t |$, contrariamente a lo que ocurre con la función $\langle \xi' | Pt \rangle$, cuya variación temporal es debida al factor de la derecha $| Pt \rangle$.

Poniendo $| a \rangle = | \xi'' \rangle$ en (47) resulta

$$i\hbar \frac{d}{dt} \langle \xi'_t | \xi'' \rangle = \int \langle \xi'_t | H_t | \xi'''_t \rangle \, d\xi'''_t \, \langle \xi'''_t | \xi'' \rangle, \tag{49}$$

que demuestra que la función de transformación $\langle \xi'_t | \xi'' \rangle$ verifica la ecuación de onda de Schrödinger. Pero $\xi_{t_0} = \xi$, y por tanto

$$\langle \xi'_{t_0} | \xi'' \rangle = \delta(\xi'_{t_0} - \xi''), \tag{50}$$

donde, igual que en el segundo miembro de (34) de § 16 con las variables $\xi_{v+1}, \ldots, \xi_u$, la función δ debe entenderse como el producto de un conjunto de factores, uno para cada variable ξ. Luego, la función de transformación $\langle \xi'_t | \xi'' \rangle$ es aquella solución de la ecuación de onda de Schrödinger para la cual en el instante t_0 las ξ tienen con certeza los valores ξ''. El cuadrado de su módulo, $| \langle \xi'_t | \xi'' \rangle |^2$, es la probabilidad relativa de que las ξ tengan los valores ξ'_t en el instante $t > t_0$ si en el instante t_0 tenían con certeza los valores ξ''. Podemos escribir $\langle \xi'_t | \xi''_{t_0} \rangle$ en lugar de $\langle \xi'_t | \xi'' \rangle$ y considerar que dependa no sólo de t sino también de t_0. Para hallar la dependencia de t_0 tomamos la ecuación compleja conjugada de (49) intercambiando t y t_0 y las primas con las segundas. Así resulta

$$-i\hbar \frac{d}{dt_0} \langle \xi'_t | \xi''_{t_0} \rangle = \int \langle \xi'_t | \xi'''_{t_0} \rangle \, d\xi'''_{t_0} \, \langle \xi'''_{t_0} | H_{t_0} | \xi''_{t_0} \rangle. \tag{51}$$

La discusión precedente relativa a la función de transformación $\langle \xi'_t | \xi'' \rangle$ es válida para cualquier conjunto completo de observables que conmuten ξ. Hemos escrito las ecuaciones para el caso en que las ξ tengan autovalores continuos, pero haciendo los cambios formales necesarios, dichas ecuaciones seguirán siendo válidas para el caso de que cualquier subconjunto de las ξ tenga autovalores discretos. Consideremos ahora un

sistema dinámico que tenga un análogo clásico y tomemos como ξ las coordenadas q. Hagamos

$$\langle q'_t | q'' \rangle = e^{iS/\hbar} \tag{52}$$

que nos define la función S de las variables q'_t, q''. Dicha función depende también explícitamente de t. (52) es una solución de la ecuación de onda de Schrödinger y si $\hbar$ se puede considerar pequeña, podemos tratarla del mismo modo que hicimos con (35). La S que figura en (52) es distinta de la S de (35), pues en (52) no figura la A, y en consecuencia, la S de (52) será compleja; pero la parte real de S será igual a la S de (35), y la parte imaginaria será del orden de $\hbar$. Luego, en el límite, cuando $\hbar \to 0$, la S que figura en (52) será igual a la de (35), y en consecuencia habrá de verificar la ecuación correspondiente a (38).

$$-\partial S/\partial t = H_c(q'_{rt}, p'_{rt}), \tag{53}$$

donde

$$p'_{rt} = \partial S/\partial q'_{rt}, \tag{54}$$

y siendo H_c el hamiltoniano del análogo clásico de nuestro sistema dinámico cuántico. Pero (52) también es solución de (51) puesta en función de las q en vez de las ξ, ecuación que es compleja conjugada de la ecuación de onda de Schrödinger en las variables q'' o q''_{t_0}. Por ello S habrá de verificar † también

$$\partial S/\partial t_0 = H_c(q''_r, p''_r), \tag{55}$$

donde

$$p''_r = -\partial S/\partial q''_r \tag{56}$$

La solución de las ecuaciones de Hamilton-Jacobi (53), (55) es la función de acción de la mecánica clásica para el intervalo t_0 a t, es decir, la integral respecto al tiempo del lagrangiano L,

$$S = \int_{t_0}^{t} L(t')\, dt' \tag{57}$$

Luego, la S definida por (52) *es el análogo cuántico de la función de acción clásica, y en el límite $\hbar \to 0$, es igual a ella.* Para hallar el análogo cuántico del lagrangiano clásico, consideremos un intervalo de tiempo infinitesimal, y pongamos $t = t_0 + \delta t$, de donde resulta que $\langle q'_{t_0+\delta t_0} | q''_{t_0} \rangle$ es el análogo de $e^{iL(t_0)\delta t/\hbar}$. Para los fines de la analogía, en lugar de considerar $L(t_0)$ como una función de las coordenadas y de las velocidades en el

---

† Para una comparación más cuidadosa de las funciones de transformación con la mecánica clásica, véase VAN VLECK, *Proc. Nat. Acad.*, 14, 178.

instante t_0, como se suele hacer generalmente, debería considerarse como una función de las coordenadas q' en el instante $t_0 + \delta t$ y de las coordenadas q'' en el instante t_0.

El principio de mínima acción de la mecánica clásica dice que la función de acción (57) es estacionaria para pequeñas variaciones de la trayectoria del sistema que no cambian los puntos extremos, es decir, para pequeñas variaciones de las q en todos los instantes intermedios entre t_0 y t que dejen fijas q_{t_0} y q_t. Analicemos qué es lo que corresponde a dicho principio en la teoría cuántica.

Pongamos

$$\exp\left\{ i \int_{t_a}^{t_b} L(t)\, dt/\hbar \right\} = \exp\{iS(t_b, t_a)/\hbar\} = B(t_b, t_a), \tag{58}$$

donde $B(t_b, t_a)$ corresponderá a $\langle q'_{t_b} | q'_{t_a} \rangle$ de la teoría cuántica. (Aquí q'_{t_a} y q'_{t_b} pueden designar autovalores de q_{t_a} y q_{t_b} distintos, pues así evitamos tener que introducir un gran número de primas en el análisis.) Consideremos el intervalo de tiempo $t_0 \to t$ dividido en un gran número de pequeños intervalos $t_0 \to t_1$, $t_1 \to t_2$, ..., $t_{m-1} \to t_m$, $t_m \to t$, que conseguimos introduciendo una sucesión de instantes intermedios t_1, t_2, ..., t_m. Entonces

$$B(t, t_0) = B(t, t_m)\, B(t_m, t_{m-1}) \ldots B(t_2, t_1)\, B(t_1, t_0). \tag{59}$$

La ecuación cuántica que corresponde a ésta, y que se deduce de la propiedad (35) de § 16 de los vectores básicos, es

$$\langle q'_t | q'_0 \rangle = \int\int \ldots \int \langle q'_t | q'_m \rangle\, dq'_m \langle q'_m | q'_{m-1} \rangle\, dq'_{m-1} \ldots$$
$$\ldots \langle q'_2 | q'_1 \rangle\, dq'_1 \langle q'_1 | q'_0 \rangle, \tag{60}$$

donde para abreviar hemos escrito q'_k en lugar de q'_{t_k}. A primera vista no parece que exista ninguna correspondencia estrecha entre (59) y (60). Sin embargo, analicemos con más cuidado el significado de (59). Cada factor B es una función de las q correspondientes a los dos extremos del intervalo a que hace referencia. Con ello, el segundo miembro de (59) es una función no sólo de q_t y de q_{t_0}, sino también de todas las q intermedias. La ecuación (59) es válida únicamente si sustituimos las q intermedias que figuran en el segundo miembro por sus valores en la trayectoria real, si bien pequeñas variaciones en ellas dejan estacionaria a S, y en consecuencia, según (58), también a $B(t, t_0)$. Lo que corresponde a las integraciones para todos los valores de las q' intermedias de (60) es el proceso de sustituir estos valores en las q intermedias. Así pues, el análogo cuántico del principio de acción queda incluido en la ley de composición (60), y la exigencia clásica

de que los valores de las q intermedias hagan a S estacionaria corresponde a la condición en mecánica cuántica de que todos los valores intermedios de las q' tienen importancia en relación con su contribución a la integral (60).

Veamos en qué forma (59) puede ser considerada como un caso límite de (60) cuando $\hbar$ es pequeño. Hemos de suponer que el integrando de (60) es de la forma $e^{iF/\hbar}$, donde F es una función de $q'_0, q'_1, q'_2, \ldots, q'_m, q'_t$ que cuando $\hbar$ tiende a cero sigue siendo continua, y por tanto cuando $\hbar$ es pequeño, el integrando es una función que oscila rápidamente. La integral de una función que oscila rápidamente será muy pequeña, salvo las contribuciones que provengan de una región del dominio de integración donde para variaciones comparativamente grandes de las q'_k F varíe relativamente poco. Tal región tiene que ser el entorno de un punto en el cual F sea estacionaria para pequeñas variaciones de q'_k. Luego, la integral (60) está determinada principalmente por el valor del integrando en un punto en el que el integrando es estacionario para pequeñas variaciones de las q' intermedias, y en consecuencia (60) se transforma en (59).

Las ecuaciones (54) y (56) expresan que las variables q'_t, p'_t están ligadas a las q'', p'' por una transformación de contacto, y constituyen una de las formas características de escribir las ecuaciones de una transformación de contacto. Existe una forma análoga de escribir las ecuaciones de una transformación unitaria en mecánica cuántica. De (52), con ayuda de (45) de § 22, obtenemos

$$\langle q'_t|p_{rt}|q''\rangle = -i\hbar\frac{\partial}{\partial q_{rt}}\langle q'_t|q''\rangle = \frac{\partial S(q'_t, q'')}{\partial q'_{rt}}\langle q'_t|q''\rangle. \tag{61}$$

Análogamente, con ayuda de (46) de § 22, resulta

$$\langle q'_t|p_r|q''\rangle = i\hbar\frac{\partial}{\partial q''_r}\langle q'_t|q''\rangle = -\frac{\partial S(q'_t, q'')}{\partial q''_r}\langle q'_t|q''\rangle. \tag{62}$$

Según la definición general de función de un conjunto de observables que conmutan, tenemos

$$\langle q'_t|f(q_t)g(q)|q''\rangle = f(q'_t)g(q'')\langle q'_t|q''\rangle \tag{63}$$

donde $f(q_t)$ y $g(q)$ son funciones respectivas de las q_t y de las q. Sea $G(q_t, q)$ cualquier función de las q_t y de las q de la forma de suma o integral de términos del tipo $f(q_t)g(q)$, de modo que todas las q_t que figuran en G estén a la izquierda de las q. Las funciones que satisfagan esta condición diremos que están *bien ordenadas*. Aplicando (63 a cada uno de los términos de G y sumando o integrando, obtenemos

$$\langle q'_t|G(q_t, q)|q''\rangle = G(q'_t, q'')\langle q'_t|q''\rangle.$$

Ahora supongamos que cada una de las p_{rt} y de las p_r puede ser expresada como función bien ordenada de las q_t y las q, y escribamos dichas funciones $p_{rt}(q_t, q)$, $p_r(q_t, q)$. Si en lugar de G ponemos dichas funciones, resulta

$$\langle q'_t|p_{rt}|q''\rangle = p_{rt}(q'_t, q'')\langle q'_t|q''\rangle,$$

$$\langle q'_t|p_r|q''\rangle = p_r(q'_t, q'')\langle q'_t|q''\rangle.$$

Comparando estas ecuaciones con (61) y (62) respectivamente, vemos que

$$p_{rt}(q'_t, q'') = \frac{\partial S(q'_t, q'')}{\partial q'_{rt}}, \qquad p_r(q'_t, q'') = -\frac{\partial S(q'_t, q'')}{\partial q''_r}$$

Lo que significa que

$$p_{rt} = \frac{\partial S(q_t, q)}{\partial q_{rt}}, \qquad p_r = -\frac{\partial S(q_t, q)}{\partial q_r}, \tag{64}$$

donde los segundos miembros de (64) han de estar expresados como funciones bien ordenadas.

Estas ecuaciones son de la misma forma que (54) y (56), pero se refieren a las variables cuánticas q_t, q que no conmutan y no a las variables algebraicas ordinarias q'_t, q''. En ellas vemos hasta qué punto las condiciones para que una transformación entre variables cuánticas sea unitaria son análogas a las condiciones para que unas variables clásicas estén ligadas por una transformación de contacto. Sin embargo, la analogía no es del todo completa puesto que la S clásica tiene que ser real y, en cambio, no existe ninguna condición sencilla para obligar a que la S de (64) también lo sea.

33. *El conjunto de Gibbs*

Hasta ahora hemos supuesto siempre que nuestro sistema dinámico está en un estado definido en cada instante de tiempo, o lo que es lo mismo, que su movimiento está especificado tan exhaustivamente y con tanta precisión como ello sea posible sin infringir los principios generales de la teoría. En la teoría clásica esto significaría, desde luego, que cada una de las coordenadas y de los momentos tendría un valor determinado. Ahora bien, podemos interesarnos por un movimiento que no sea conocido con la máxima precisión posible. En esta sección consideremos los métodos que deben emplearse en este caso.

En mecánica clásica el procedimiento consiste en introducir lo que se denomina un *conjunto de Gibbs*, cuya idea es la siguiente. Consideramos

todas las coordenadas y momentos dinámicos como coordenadas cartesianas de un cierto espacio que llamamos *espacio de las fases,* cuyo número de dimensiones será dos veces el número de grados de libertad del sistema. Todo estado del sistema puede ser representado por un punto de dicho espacio. Dicho punto se moverá de acuerdo con las ecuaciones del movimiento clásicas (14). Pero supongamos ahora que no sabemos en qué estado está en cada instante de tiempo, sino solamente que está en uno u otro de un conjunto de posibles estados de acuerdo con una ley probabilística conocida. Podríamos representarlo mediante un fluido en el espacio de las fases, para el que la masa contenida en un volumen cualquiera del espacio de las fases fuera igual a la probabilidad total de que el sistema estuviera en un estado cuyo punto representativo estuviera comprendido en el interior de dicho volumen. Cada partícula de fluido se movería de acuerdo con las ecuaciones de movimiento (14). Si introducimos la densidad p del fluido en cada punto como la probabilidad por unidad de volumen del espacio de las fases de que el sistema esté en un estado cuyo punto representativo está en el entorno del punto considerado, obtendremos la ecuación de conservación

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\sum_r \left\{ \frac{\partial}{\partial q_r}\left(\rho \frac{dq_r}{dt}\right) + \frac{\partial}{\partial p_r}\left(\rho \frac{dp_r}{dt}\right)\right\} \\ &= -\sum_r \left\{ \frac{\partial}{\partial q_r}\left(\rho \frac{\partial H}{\partial p_r}\right) - \frac{\partial}{\partial p_r}\left(\rho \frac{\partial H}{\partial q_r}\right)\right\} \\ &= -[\rho, H]. \end{aligned} \tag{65}$$

que puede ser considerada como la ecuación de movimiento del fluido, pues determina la densidad ρ en cualquier instante si conocemos la densidad ρ inicialmente como función de las q y de las p. Salvo el signo menos, esta ecuación es de la misma forma que la ecuación de movimiento ordinaria (15) para una variable dinámica.

Al exigir que la probabilidad total de que el sistema esté en algún estado sea igual a uno obtenemos una condición de normalización para ρ

$$\int\int \rho \; dqdp = 1, \tag{66}$$

donde la integración está extendida a todo el espacio de las fases y donde dq y dp simbolizan el producto de todos los dq y de todos los dp. Si β es una función cualquiera de las variables dinámicas, el valor medio de β será

$$\int\int \beta\rho \; dqdp. \tag{67}$$

Si en lugar de la densidad considerada tomamos una ρ que difiera de la anterior en un factor k positivo y constante, la teoría no cambia más que de

forma trivial, pero en cambio nos facilita la interpretación. En lugar de (66) tendremos ahora

$$\iint \rho \, dq dp = k.$$

Con esta nueva densidad la imagen del fluido puede ser un conjunto de k sistemas dinámicos similares, cada uno de los cuales se mueve independientemente en el mismo espacio, sin ninguna perturbación o interacción mutua. La densidad en cualquier punto será igual al número probable o promedio de sistemas en el entorno de cualquier estado por unidad de volumen del espacio de las fases, y la expresión (67) representará el valor medio total de β para todos los sistemas. Tal conjunto de sistemas dinámicos, que constituye el llamado conjunto de Gibbs, generalmente sólo se puede realizar en la práctica como aproximación grosera, pero de todos modos constituye una abstracción teórica muy útil.

Veremos que existe una densidad ρ correspondiente en mecánica cuántica, que tiene propiedades análogas a la anterior. Fue introducida por primera vez por von Neumann. Su existencia es bastante sorprendente si se tiene en cuenta que el espacio de las fases no tiene ningún significado en mecánica cuántica, ya que no hay ningún modo de asignar simultáneamente valores numéricos a las q y a las p.

Consideremos un sistema dinámico que en un cierto instante esté en uno u otro de un conjunto de posibles estados de acuerdo con una ley probabilística conocida. Dichos estados pueden constituir un conjunto discreto o bien un conjunto continuo o ambas cosas a la vez. Para concretar consideraremos el caso de un conjunto discreto que supondremos numerado por un parámetro m. Sean $|m\rangle$ los kets normalizados correspondientes a dichos estados, y P_m la probabilidad de que el sistema esté en el estado m-ésimo. Definimos entonces la densidad cuántica ρ por

$$\rho = \sum_m |m\rangle P_m \langle m|. \tag{68}$$

Sea ρ' un autovalor cualquiera de ρ y $|\rho'\rangle$ un autoket perteneciente a dicho autovalor. En ese caso

$$\sum_m |m\rangle P_m \langle m|\rho'\rangle = \rho|\rho'\rangle = \rho'|\rho'\rangle$$

y por tanto

$$\sum_m \langle\rho'|m\rangle P_m \langle m|\rho'\rangle = \rho'\langle\rho'|\rho'\rangle$$

o sea

$$\sum_m P_m |\langle m|\rho'\rangle|^2 = \rho'\langle\rho'|\rho'\rangle.$$

Pero como P_m es una probabilidad no puede ser negativa. Por tanto, ρ' no puede ser negativo. Luego ρ no tiene ningún autovalor negativo, análogamente a lo que ocurre con la densidad clásica ρ que no es nunca negativa.

Hallemos la ecuación de movimiento para la densidad cuántica ρ. En la imagen de Schrödinger los kets y bras de (68) variarán con el tiempo de acuerdo con la ecuación de Schrödinger (5) y su conjugada imaginaria, en tanto que las P_m permanecerán constantes, pues mientras el sistema permanezca inalterado no puede pasar de un estado correspondiente a un ket que verifica la ecuación de Schrödinger a otro estado correspondiente a otro. En consecuencia,

$$\begin{aligned} i\hbar \frac{d\rho}{dt} &= \sum_m i\hbar \left\{ \frac{d|m\rangle}{dt} P_m \langle m| + |m\rangle P_m \frac{d\langle m|}{dt} \right\} \\ &= \sum_m \{H|m\rangle P_m \langle m| - |m\rangle P_m \langle m|H\} \\ &= H\rho - \rho H. \end{aligned} \tag{69}$$

Ésta es la análoga cuántica de la ecuación del movimiento clásica (65). La ρ cuántica, igual que la clásica, está determinada en todo instante si se conoce su valor inicial.

Según la hipótesis de § 12, el valor esperado de cualquier observable β cuando el sistema está en el estado m es $\langle m|\beta|m\rangle$. Por tanto, si P_m es la ley de probabilidad de que el sistema esté en los distintos estados m, el valor esperado de β será $\sum_m P_m \langle m|\beta|m\rangle$. Introduciendo una representación que tenga un conjunto de kets básicos $|\xi'\rangle$ discreto, esto equivale a

$$\begin{aligned} \sum_{m\xi'} P_m \langle m|\xi'\rangle\langle\xi'|\beta|m\rangle &= \sum_{\xi' m} \langle\xi'|\beta|m\rangle P_m \langle m|\xi'\rangle \\ &= \sum_{\xi'} \langle\xi'|\beta\rho|\xi'\rangle = \sum_{\xi'} \langle\xi'|\rho\beta|\xi'\rangle, \end{aligned} \tag{70}$$

siendo el último paso fácilmente comprobable con la ley de multiplicación de matrices, ecuación (44) de § 17. Las expresiones (70) son las análogas de la (67) de la teoría clásica. Allí donde en mecánica clásica hayamos de multiplicar β por ρ e integrar dicho producto para todo el espacio de las fases, en la teoría cuántica habremos de multiplicar β por ρ en cualquier orden, y tomar la suma de los elementos diagonales † de dicho producto en una representación. Si en la representación existe un conjunto continuo de vectores básicos $|\xi'\rangle$, en lugar de (70) tendremos

$$\int \langle\xi'|\beta\rho|\xi'\rangle \, d\xi' = \int \langle\xi'|\rho\beta|\xi'\rangle \, d\xi', \tag{71}$$

y entonces en lugar de sumar los elementos diagonales habremos de reali-

† *N. del T.:* La suma de todos los elementos diagonales de una matriz se denomina *traza*.

zar un proceso de 'integración a lo largo de la diagonal'. Definiremos la suma diagonal de $\beta\rho$ en el caso continuo por (71). Se puede comprobar fácilmente a partir de las propiedades de las funciones de transformación (56) de § 18, que la suma de los elementos diagonales es igual en todas las representaciones.

De la condición de que los $|m\rangle$ estén normalizados se deduce, para el caso de ξ' discretos

$$\sum_{\xi'} \langle\xi'|\rho|\xi'\rangle = \sum_{\xi' m} \langle\xi'|m\rangle P_m \langle m|\xi'\rangle = \sum_m P_m = 1, \tag{72}$$

pues la probabilidad de que el sistema esté en algún estado es igual a uno. Ésta es la ecuación análoga de la (66). La probabilidad de que el sistema esté en el estado ξ', o la probabilidad de que los observables ξ que son diagonales en la representación tengan los valores ξ' es, según la interpretación de los representantes de un ket (51) de § 18,

$$\sum_m |\langle\xi'|m\rangle|^2 P_m = \langle\xi'|\rho|\xi'\rangle, \tag{73}$$

que da significado a cada uno de los términos del primer miembro de (72). Para ξ' continuos, el segundo miembro de (73) da la probabilidad de que los ξ tengan valores en el entorno de ξ' por unidad de intervalo de variación de las ξ'.

Igual que en la teoría clásica, podemos considerar una densidad igual a k veces la ρ que hemos definido y decir que representa un conjunto de Gibbs constituido por k sistemas dinámicos iguales, entre los que no hay perturbación o interacción alguna. Así el último miembro de (72) será igual a k, y (70) o (71) nos darán el valor esperado de β para todos los elementos del conjunto, mientras que (73) representará la probabilidad total de que un elemento del conjunto tenga unos valores de las ξ iguales a ξ', o en el entorno de ξ' por unidad de intervalo de variación de los valores ξ'.

Una aplicación importante del conjunto de Gibbs es su aplicación a un sistema dinámico en equilibrio termodinámico cuyo contorno está a una temperatura dada T. Gibbs demostró que dicho sistema está caracterizado en mecánica clásica por una densidad

$$\rho = ce^{-H/kT}, \tag{74}$$

donde H es el hamiltoniano, que en este caso es independiente del tiempo, k es la constante de Boltzmann, y c es un número que se elige de modo que se verifique la condición de normalización (66). Esta fórmula puede ser tomada a su vez en mecánica cuántica sin ninguna modificación. Para altas temperaturas, (74) se reduce a $\rho = c$, que sustituido en el segundo miembro de (73) de $c\langle\xi'|\xi'\rangle = c$, en el caso de ξ' discretos. Vemos pues, que *a altas temperaturas todos los estados discretos son igualmente probables.*

VI

APLICACIONES ELEMENTALES

34. *El oscilador armónico*

Un ejemplo sencillo e interesante de un sistema dinámico en mecánica cuántica lo constituye el oscilador armónico. Tiene importancia en la teoría general pues es un elemento básico en la teoría de la radiación. Las únicas variables dinámicas que se necesitan para describir dicho sistema son una coordenada q y su momento canónico conjugado p. En mecánica clásica su hamiltoniano es

$$H = \frac{1}{2m}(p^2 + m^2\omega^2 q^2), \tag{1}$$

donde m es la masa de la partícula que oscila, y ω es igual a 2π veces la frecuencia. Supondremos que en mecánica cuántica es válido el mismo hamiltoniano. Este hamiltoniano junto con la condición cuántica (10) de § 22 definen completamente el sistema.

Las ecuaciones del movimiento de Heisenberg son

$$\left.\begin{aligned} \dot{q}_t &= [q_t, H] = p_t/m, \\ \dot{p}_t &= [p_t, H] = -m\omega^2 q_t. \end{aligned}\right\} \tag{2}$$

Es conveniente introducir la variable dinámica compleja sin dimensiones

$$\eta = (2m\hbar\omega)^{-\frac{1}{2}}(p + im\omega q). \tag{3}$$

Las ecuaciones del movimiento (2) dan

$$\dot{\eta}_t = (2m\hbar\omega)^{-\frac{1}{2}}(-m\omega^2 q_t + i\omega p_t) = i\omega\eta_t.$$

Esta ecuación se puede integrar y resulta

$$\eta_t = \eta_0\, e^{i\omega t}, \tag{4}$$

donde η_0 es un operador lineal independiente de t, y es igual al valor de η_t en el instante inicial $t = 0$. Las ecuaciones anteriores son idénticas a las de la teoría clásica.

Podemos expresar q y p en función de η y de su compleja conjugada $\bar{\eta}$, y poner todo en función exclusivamente de las variables η y $\bar{\eta}$. Tenemos

$$\begin{aligned} \hbar\omega\eta\bar{\eta} &= (2m)^{-1}(p+im\omega q)(p-im\omega q) \\ &= (2m)^{-1}[p^2+m^2\omega^2q^2+im\omega(qp-pq)] \\ &= H-\tfrac{1}{2}\hbar\omega \end{aligned} \tag{5}$$

y análogamente

$$\hbar\omega\bar{\eta}\eta = H+\tfrac{1}{2}\hbar\omega. \tag{6}$$

Luego,

$$\bar{\eta}\eta-\eta\bar{\eta} = 1. \tag{7}$$

La ecuación (5) o (6) expresan H en función de η y $\bar{\eta}$, y (7) da la relación de conmutación entre η y $\bar{\eta}$. De (5) resulta

$$\hbar\omega\bar{\eta}\eta\bar{\eta} = \bar{\eta}H-\tfrac{1}{2}\hbar\omega\bar{\eta}$$

y de (6)

$$\hbar\omega\bar{\eta}\eta\bar{\eta} = H\bar{\eta}+\tfrac{1}{2}\hbar\omega\bar{\eta}.$$

Luego,

$$\bar{\eta}H-H\bar{\eta} = \hbar\omega\bar{\eta}. \tag{8}$$

Asimismo, la (7) conduce por inducción a

$$\bar{\eta}\eta^n-\eta^n\bar{\eta} = n\eta^{n-1} \tag{9}$$

para todo valor de n entero y positivo, pues multiplicando (9) por η a la izquierda se obtiene (9) con $n+1$ en lugar de n.

Sea H' un autovalor de H y sea $|H'\rangle$ un autoket perteneciente a él. Según (5)

$$\hbar\omega\langle H'|\eta\bar{\eta}|H'\rangle = \langle H'|H-\tfrac{1}{2}\hbar\omega|H'\rangle = (H'-\tfrac{1}{2}\hbar\omega)\langle H'|H'\rangle,$$

Pero $\langle H'|\eta\bar{\eta}|H'\rangle$ es el cuadrado de la longitud del ket $\bar{\eta}|H'\rangle$, y por tanto,

$$\langle H'|\eta\bar{\eta}|H'\rangle \geqslant 0,$$

siendo válido el signo igual sólo cuando $\bar{\eta}|H'\rangle = 0$. Asimismo, $\langle H'|H'\rangle > 0$. Luego,

$$H' \geqslant \tfrac{1}{2}\hbar\omega, \tag{10}$$

siendo válido el signo igual sólo cuando $\bar{\eta}|H'\rangle = 0$. Como H tiene la forma de suma de cuadrados (1), es de esperar que sus autovalores sean todos positivos o nulos (pues el valor medio de H para cualquier estado tiene que ser positivo o nulo). Pero además ahora hemos obtenido una condición más restringida (10).

De (8) resulta

$$H\bar{\eta}|H'\rangle = (\bar{\eta}H-\hbar\omega\bar{\eta})|H'\rangle = (H'-\hbar\omega)\bar{\eta}|H'\rangle. \tag{11}$$

Ahora bien, si $H' \neq \frac{1}{2}\hbar\omega$, $\bar{\eta}|H'\rangle$ es distinto de cero, y según (11) es un

autoket de H perteneciente al autovalor $H' - \hbar\omega$. Luego, si H' es cualquier autovalor de H distinto de $\frac{1}{2}\hbar\omega$, entonces $H' - \hbar\omega$ es también autovalor de H. Podemos repetir el razonamiento y deducir que si $H' - \hbar\omega \neq \frac{1}{2}\hbar\omega$, entonces también $H' - 2\hbar\omega$ es autovalor de H. Reiterando el procedimiento obtenemos la serie de autovalores H', $H' - \hbar\omega$, $H' - 2\hbar\omega$, $H' - 3\hbar\omega$, ..., que, como en virtud de (10) no puede extenderse indefinidamente, ha de terminar forzosamente con el valor $\frac{1}{2}\hbar\omega$. Asimismo de la ecuación compleja conjugada de la (8) resulta

$$H\eta|H'\rangle = (\eta H + \hbar\omega\eta)|H'\rangle = (H' + \hbar\omega)\eta|H'\rangle,$$

que muestra que salvo en el caso de que $\dot{\eta}|H'\rangle = 0$, $H' + \hbar\omega$ es otro autovalor de H uno de cuyos autokets es $\dot{\eta}|H'\rangle$. La posibilidad de que $\bar{\eta}|H'\rangle = 0$ puede ser desechada, pues nos conduciría a

$$0 = \hbar\omega\bar{\eta}\eta|H'\rangle = (H + \tfrac{1}{2}\hbar\omega)|H'\rangle = (H' + \tfrac{1}{2}\hbar\omega)|H'\rangle,$$

en contradicción con (10). Luego, $H' + \hbar\omega$ siempre es otro autovalor de H y por tanto también lo son $H' + 2\hbar\omega$, $H' + 3\hbar\omega$, etc. Por consiguiente, los autovalores de H son la sucesión de números

$$\tfrac{1}{2}\hbar\omega, \qquad \tfrac{3}{2}\hbar\omega, \qquad \tfrac{5}{2}\hbar\omega, \qquad \tfrac{7}{2}\hbar\omega, \ \ldots \tag{12}$$

que se extiende hasta infinito. Éstos son los valores posibles de la energía para el oscilador armónico.

Sea $|0\rangle$ un autoket de H perteneciente al mínimo autovalor $\frac{1}{2}\hbar\omega$; por tanto,

$$\bar{\eta}|0\rangle = 0, \tag{13}$$

y formemos la sucesión de kets

$$|0\rangle, \qquad \eta|0\rangle, \qquad \eta^2|0\rangle, \qquad \eta^3|0\rangle, \qquad \ldots . \tag{14}$$

Todos ellos serán autokets de H pertenecientes respectivamente a la sucesión de autovalores (12). De (9) y (13) se deduce

$$\bar{\eta}\eta^n|0\rangle = n\eta^{n-1}|0\rangle \tag{15}$$

para todo entero n no negativo. Luego, el conjunto de kets (14) es tal que η o $\bar{\eta}$ aplicado a cualquier ket del conjunto da un ket dependiente del conjunto. Ahora bien, como todas las variables dinámicas de nuestro problema se puede expresar en función de η y $\bar{\eta}$, los kets (14) constituyen un conjunto completo (de no ser así, existirían otras variables dinámicas). Hay un solo ket del conjunto para cada autovalor (12) de H, luego, H constituye por sí sólo un conjunto completo de observables. Los kets (14) corresponden a los distintos estados estacionarios del oscilador. El estado estacionario con energía $(n + \frac{1}{2})\hbar\omega$, correspondiente a $\eta^n|0\rangle$, se denomina estado cuántico de orden n.

El cuadrado de la longitud del ket $\dot{\eta}^n|0\rangle$ según (15) es

$$\langle 0|\bar{\eta}^n\bar{\eta}^n|0\rangle = n\langle 0|\bar{\eta}^{n-1}\eta^{n-1}|0\rangle$$

Por inducción obtenemos

$$\langle 0|\bar{\eta}^n\eta^n|0\rangle = n! \tag{16}$$

si $|0\rangle$ estaba normalizado. Por tanto, los kets (14) multiplicados respectivamente por los coeficientes $n!^{-\frac{1}{2}}$, con $n = 0, 1, 2, \ldots$, constituyen los kets básicos de una representación, a saber, la representación en la que H es diagonal. Todo ket $|x\rangle$ se puede desarrollar en la forma

$$|x\rangle = \sum_0^\infty x_n\, \eta^n|0\rangle, \tag{17}$$

siendo los x_n números. Por este procedimiento el ket $|x\rangle$ está en correspondencia con una serie de potencias $\sum x_n\, \dot{\eta}^n$ de la variable $\dot{\eta}$, donde los distintos términos corresponden a los distintos estados estacionarios. Si $|x\rangle$ está normalizado, define un estado para el que la probabilidad de que el oscilador armónico esté en el estado cuántico de orden n, o sea, de que H tenga el valor $(n + \frac{1}{2})\hbar\omega$, es

$$P_n = n!\ |x_n|^2, \tag{18}$$

según se deduce del mismo razonamiento que nos condujo a la (51) de § 18.

El ket $|0\rangle$ puede ser considerado como un ket standard y la serie de potencias de $\dot{\eta}$ como una función de onda, pues todo ket se puede expresar como dicha función de onda multiplicada por el ket standard. Así obtenemos un tipo de función de onda que difiere de las que utilizamos corrientemente, y que introdujimos en (62) de § 20, en que es función de una variable dinámica compleja $\dot{\eta}$ y no de un observable. Esta representación la utilizó por primera vez por V. Fock, y la llamaremos representación de Fock. En muchas aplicaciones es la representación más conveniente para describir estados del oscilador armónico. El ket standard $|0\rangle$ verifica la condición (13) que sustituye a las condiciones (43) de § 22 del ket standard de la representación de Schrödinger.

Introduzcamos la representación de Schrödinger con q diagonal para obtener los representantes de los estados estacionarios. De (13) y (3) resulta

$$(p - im\omega q)|0\rangle = 0,$$

de modo que

$$\langle q'|p - im\omega q|0\rangle = 0.$$

Con ayuda de (45) de § 22, ésta equivale a

$$\hbar\frac{\partial}{\partial q'}\langle q'|0\rangle + m\omega q'\langle q'|0\rangle = 0. \tag{19}$$

La solución de esta ecuación diferencial es

$$\langle q'|0\rangle = (m\omega/\pi\hbar)^{\frac{1}{4}} e^{-m\omega q'^2/2\hbar}, \tag{20}$$

donde hemos tomado el coeficiente numérico de modo que $|0\rangle$ esté normalizado. Éste es el representante del *estado fundamental,* nombre que se da al estado de menor energía. Los representantes de los demás estados estacionarios se pueden deducir a partir de éste. Según (3) tenemos

$$\begin{aligned}\langle q'|\eta^n|0\rangle &= (2m\hbar\omega)^{-n/2}\langle q'|(p+im\omega q)^n|0\rangle \\ &= (2m\hbar\omega)^{-n/2} i^n \left(-\hbar\frac{\partial}{\partial q'} + m\omega q'\right)^n \langle q'|0\rangle \\ &= i^n (2m\hbar\omega)^{-n/2} (m\omega/\pi\hbar)^{\frac{1}{4}} \left(-\hbar\frac{\partial}{\partial q'} + m\omega q'\right)^n e^{-m\omega q'^2/2\hbar}. \end{aligned} \tag{21}$$

Esta expresión se puede desarrollar fácilmente para pequeños valores de n. El resultado es de la forma $e^{-m\omega q'^2/2\hbar}$ multiplicado por un polinomio de grado n en q'. Para que el representante del estado cuántico de orden n esté normalizado tenemos que agregar en (21) el factor $n!^{-\frac{1}{2}}$. Los factores de fase i^n se pueden suprimir.

35. *Momento angular*

Consideremos una partícula caracterizada por sus tres coordenadas cartesianas x, y, z y sus tres momentos conjugados p_x, p_y, p_z. El momento angular respecto al origen se define, igual que en la teoría clásica, por

$$m_x = yp_z - zp_y \qquad m_y = zp_x - xp_z \qquad m_z = xp_y - yp_x, \tag{22}$$

o bien por la ecuación vectorial

$$\mathbf{m} = \mathbf{x} \times \mathbf{p}.$$

Hemos de calcular los P. B. de las componentes del momento angular con las variables dinámicas x, p_x, etc., y consigo mismas. El modo más fácil de hacerlo es utilizar las reglas (4) y (5) de § 21. Así

$$\left.\begin{aligned}[m_z, x] &= [xp_y - yp_x, x] = -y[p_x, x] = y, \\ [m_z, y] &= [xp_y - yp_x, y] = x[p_y, y] = -x,\end{aligned}\right\} \tag{23}$$

$$[m_z, z] = [xp_y - yp_x, z] = 0, \tag{24}$$

y análogamente

$$[m_z, p_x] = p_y, \qquad [m_z, p_y] = -p_x, \tag{25}$$

$$[m_z, p_z] = 0, \tag{26}$$

con las relaciones correspondientes para m_x y m_y. Tenemos además

$$\left.\begin{aligned} [m_x, m_z] &= [zp_x - xp_z, m_z] = z[p_x, m_z] - [x, m_z]p_z \\ &= -zp_y + yp_z = m_x, \\ [m_z, m_x] &= m_y, \qquad [m_x, m_y] = m_z. \end{aligned}\right\} \quad (27)$$

Estos resultados son los mismos que en la teoría clásica. Los signos de (23), (25) y (27) son fáciles de recordar por la regla siguiente: cuando las tres variables dinámicas, es decir, las dos del P. B. de la izquierda y la del resultado de la derecha están en permutación circular de (xyz), el signo es positivo, y en caso contrario es negativo. Las ecuaciones (27) se pueden escribir en notación vectorial

$$\mathbf{m} \times \mathbf{m} = i\hbar\mathbf{m}. \quad (28)$$

Supongamos ahora que existan varias partículas de momentos angulares $\mathbf{m}_1$, $\mathbf{m}_2$, ... Cada uno de estos momentos angulares vectoriales verificará la relación (28), o sea,

$$\mathbf{m}_r \times \mathbf{m}_r = i\hbar\mathbf{m}_r,$$

y además cada uno de ellos conmutará con todos los demás, y en consecuencia

$$\mathbf{m}_r \times \mathbf{m}_s + \mathbf{m}_s \times \mathbf{m}_r = 0 \quad (r \neq s).$$

Luego si $\mathbf{M} = \sum_r \mathbf{m}_r$ es el momento angular total,

$$\begin{aligned} \mathbf{M} \times \mathbf{M} &= \sum_{rs} \mathbf{m}_r \times \mathbf{m}_s = \sum_r \mathbf{m}_r \times \mathbf{m}_r + \sum_{r<s} (\mathbf{m}_r \times \mathbf{m}_s + \mathbf{m}_s \times \mathbf{m}_r) \\ &= i\hbar \sum_r \mathbf{m}_r = i\hbar\mathbf{M}. \end{aligned} \quad (29)$$

Este resultado es de la misma forma que (28), y por tanto, las componentes del momento angular total $\mathbf{M}$ de cualquier número de partículas verifican las mismas relaciones de conmutación que el momento angular de una sola partícula.

Sean A_x, A_y, A_z las tres coordenadas o las tres componentes del momento de una de las partículas. Las A conmutarán con los momentos angulares de las otras partículas, luego según (23), (24), (25) y (26) será

$$[M_z, A_x] = A_y, \qquad [M_z, A_y] = -A_x, \qquad [M_z, A_z] = 0. \quad (30)$$

Si B_x, B_y, B_z es un segundo conjunto de tres cantidades que representan las tres coordenadas o las tres componentes del momento de una de las partículas, también verificarán relaciones similares a (30). Tendremos entonces

$$\begin{aligned} &[M_z, A_x B_x + A_y B_y + A_z B_z] \\ &\quad = [M_z, A_x]B_x + A_x[M_z, B_x] + [M_z, A_y]B_y + A_y[M_z B_y] \\ &\quad = A_y B_x + A_x B_y - A_x B_y - A_y B_x \\ &\quad = 0. \end{aligned}$$

Por tanto, el producto escalar $A_x B_x + A_y B_y + A_z B_z$ conmuta con M_z, y análogamente también con M_x y M_y. Consideremos el producto vectorial

$$\mathbf{A} \times \mathbf{B} = \mathbf{C}$$

o sea,

$$A_y B_z - A_z B_y = C_x, \quad A_z B_x - A_x B_z = C_y, \quad A_x B_y - A_y B_x = C_z.$$

Tenemos

$$[M_z, C_x] = -A_x B_z + A_z B_x = C_y$$

y análogamente

$$[M_z, C_y] = -C_x, \qquad [M_z, C_z] = 0.$$

Estas ecuaciones son a su vez de la forma (30) con **C** en lugar de **A**. De aquí resulta que las ecuaciones (30) son válidas para las componentes de cualquier vector que podamos formar a partir de las variables dinámicas, y que todo escalar conmuta con **M**.

Podemos introducir operadores lineales de rotación R alrededor del origen análogamente a como hicimos con los operadores de traslación D en § 25. Para una rotación de ángulo $\delta\phi$ alrededor del eje z, considerando $\delta\phi$ como un infinitésimo, obtenemos el operador límite correspondiente a 64 de § 25,

$$\lim_{\delta\phi \to 0} (R - 1)/\delta\phi,$$

que llamaremos *operador de rotación* alrededor del eje z y designaremos por r_z. Igual que los operadores de traslación, r_z es también un operador lineal imaginario puro que está indeterminado en un número aditivo imaginario. En correspondencia con (66) de § 25, la variación de primer orden de $\delta\phi$ que experimenta una variable dinámica v al realizar un giro de ángulo $\delta\phi$ pequeño alrededor del eje z es

$$\delta\phi(r_z v - v r_z). \tag{31}$$

Pero las variaciones de las tres componentes A_x, A_y y A_z de un vector, debidas a un giro de todos los aparatos de medida en un ángulo $\delta\phi$ (hacia la derecha) alrededor del eje z, son respectivamente $\delta\phi A_y$, $-\delta\phi A_x$ y 0, mientras que los escalares permanecen invariantes en el giro. Igualando estos cambios a (31), resulta

$$r_z A_x - A_x r_z = A_y, \qquad r_z A_y - A_y r_z = -A_x,$$
$$r_z A_z - A_z r_z = 0,$$

y que r_z conmuta con cualquier escalar. Comparando estos resultados con (30), vemos que $i\hbar r_z$ verifica las mismas relaciones de conmutación que M_z. Luego, su diferencia $M_z - i\hbar r_z$ tiene que conmutar con todas las variables dinámicas, y por tanto, es un número. Dicho número, que necesariamente ha de ser real, pues M_z y $i\hbar r_z$ son ambos reales, puede hacerse igual a cero

eligiendo convenientemente el número imaginario puro arbitrario que se podía sumar a r_z. Entonces resultará

$$M_z = i\hbar r_z. \tag{32}$$

Para M_x y M_y tendremos las ecuaciones correspondientes. Estas ecuaciones son análogas a (69) de § 25. Luego, *el momento angular total está asociado con los operadores de rotación del mismo modo en que el momento total está asociado con los operadores de traslación.* Esta conclusión es válida tomando el origen en cualquier punto.

Los razonamientos anteriores son válidos para los momentos angulares debidos al movimiento de las partículas, que para cada partícula vienen dados por (22). En la teoría atómica interviene otra clase de momento angular, que es el *momento angular de spin.* Al primero se le llama *momento angular orbital* para diferenciarlo del segundo. El momento angular de spin debe atribuirse a algún movimiento interno de la partícula, pues está asociado a grados de libertad distintos de los que intervienen en la descripción del movimiento de la partícula como un todo, y por tanto, las variables dinámicas que representan el spin tienen que conmutar con x, y, z, p_x, p_y y p_z. Estrictamente, el spin no tiene correspondiente en mecánica clásica, y en consecuencia el método de analogía clásica no es conveniente para su estudio. Sin embargo, podemos construir una teoría del spin postulando simplemente que las componentes del momento angular de spin estén relacionadas con operadores de rotación del mismo modo que los momentos angulares orbitales. Es decir, postulando que la ecuación (32) también sea válida para M_z igual a la componente z del momento angular de spin de una partícula y r_z igual al operador de rotación alrededor del eje z relativo a los estados de spin de la partícula. Con esta hipótesis, las relaciones de conmutación de las componentes del momento angular de spin **M** con cualquier vector **A** relativo al spin han de ser de la forma (30), y en consecuencia, tomando **A** como el propio momento angular de spin se obtiene la ecuación (29) para el spin. Así pues, (29) es válida en general para cualquier suma de momentos angulares orbitales y de spin, y análogamente (30) también será válida para **M** igual al momento angular total orbital y de spin y **A** una variable dinámica vectorial cualquiera. Asimismo la relación entre momentos angulares y rotaciones tiene también validez general.

Como consecuencia inmediata de esta relación podemos deducir la *ley de conservación del momento angular.* Para un sistema aislado, el hamiltoniano debe ser invariante frente a las rotaciones en torno al origen, o lo que es lo mismo, ha de ser un escalar; por tanto, tiene que conmutar con el momento angular alrededor del origen. Luego, el momento angular es una constante del movimiento. En este razonamiento el origen puede ser cualquier punto.

Como segunda consecuencia inmediata podemos ver que *un estado con*

momento angular total nulo tiene simetría esférica. Dicho estado corresponderá a un ket $|S\rangle$ que verificará

$$M_x|S\rangle = M_y|S\rangle = M_z|S\rangle = 0,$$

y, por tanto,

$$r_x|S\rangle = r_y|S\rangle = r_z|S\rangle = 0.$$

Esto indica que el ket $|S\rangle$ es invariante frente a las rotaciones infinitesimales, y en consecuencia también frente a rotaciones finitas, ya que éstas se pueden construir a partir de las primeras. Luego, el estado tiene simetría esférica. El teorema recíproco de que *un estado con simetría esférica tiene momento angular total nulo,* también es cierto, si bien la demostración ya no es tan sencilla. Un estado con simetría esférica corresponde a un ket $|S\rangle$ cuya dirección no se modifica por una rotación. Luego, la variación de $|S\rangle$ al aplicarle un operador r_x, r_y o r_z tiene que ser un múltiplo numérico de $|S\rangle$, por ejemplo,

$$r_x|S\rangle = c_x|S\rangle, \qquad r_y|S\rangle = c_y|S\rangle, \qquad r_z|S\rangle = c_z|S\rangle,$$

donde los c son números. Así obtenemos

$$M_x|S\rangle = i\hbar c_x|S\rangle, \qquad M_y|S\rangle = i\hbar c_y|S\rangle,$$
$$M_z|S\rangle = i\hbar c_z|S\rangle, \tag{33}$$

Estas ecuaciones están en contradicción con las relaciones de conmutación (29) entre M_x, M_y y M_z salvo si $c_x = c_y = c_z = 0$, en cuyo caso el estado tiene momento angular total nulo. (33) constituye un ejemplo de autoket común a los tres operadores lineales M_x, M_y y M_z que no conmutan entre sí, y ello sólo es posible si los tres autovalores correspondientes son cero.

36. *Propiedades del momento angular*

Hay algunas propiedades del momento angular que se pueden deducir sin más de las relaciones de conmutación entre las tres componentes. Tales propiedades son válidas tanto para el momento angular orbital como para el de spin. Sean m_x, m_y y m_z las tres componentes de un momento angular, e introduzcamos la cantidad β definida por la ecuación

$$\beta = m_x^2 + m_y^2 + m_z^2.$$

Como β es un escalar ha de conmutar con m_x, m_y y m_z. Sea un sistema dinámico cuyas únicas variables dinámicas sean m_x, m_y y m_z. En ese caso β conmuta con todos los operadores y, por tanto, ha de ser un número. Podemos estudiar este sistema dinámico de forma parecida a cómo hicimos con el oscilador armónico en § 34.

Pongamos

$$m_x - im_y = \eta.$$

A partir de las relaciones de conmutación (27) deducimos

$$\bar{\eta}\eta = (m_x + im_y)(m_x - im_y) = m_x^2 + m_y^2 - i(m_x m_y - m_y m_x)$$
$$= \beta - m_z^2 + \hbar m_z \tag{34}$$

y análogamente

$$\eta\bar{\eta} = \beta - m_z^2 - \hbar m_z. \tag{35}$$

Luego, resulta

$$\bar{\eta}\eta - \eta\bar{\eta} = 2\hbar m_z. \tag{36}$$

Asimismo

$$m_z\eta - \eta m_z = i\hbar m_y - \hbar m_x = -\hbar\eta, \tag{37}$$

Supondremos que las componentes del momento angular son observables y que, por tanto, m_z tiene autovalores. Sea m'_z uno de ellos y sea $|m'_z\rangle$ un autoket perteneciente a dicho autovalor. Según (34)

$$\langle m'_z|\bar{\eta}\eta|m'_z\rangle = \langle m'_z|\beta - m_z^2 + \hbar m_z|m'_z\rangle = (\beta - m'^2_z + \hbar m'_z)\langle m'_z|m'_z\rangle.$$

El primer miembro de esta ecuación es el cuadrado de la longitud del ket $\eta|m'_z\rangle$ y, por tanto, es mayor o igual que cero. El caso de igualdad se tendrá si y sólo si $\eta|m'_z\rangle = 0$. Luego,

$$\beta - m'^2_z + \hbar m'_z \geqslant 0,$$

o bien

$$\beta + \tfrac{1}{4}\hbar^2 \geqslant (m'_z - \tfrac{1}{2}\hbar)^2. \tag{38}$$

Por tanto,

$$\beta + \tfrac{1}{4}\hbar^2 \geqslant 0$$

Si definimos el número k por

$$k + \tfrac{1}{2}\hbar = (\beta + \tfrac{1}{4}\hbar^2)^{\frac{1}{2}} = (m_x^2 + m_y^2 + m_z^2 + \tfrac{1}{4}\hbar^2)^{\frac{1}{2}}, \tag{39}$$

con lo cual $k \geqslant -\frac{1}{2}\hbar$, la desigualdad (38) se convierte en

$$k + \tfrac{1}{2}\hbar \geqslant |m'_z - \tfrac{1}{2}\hbar|$$

o bien

$$k + \hbar \geqslant m'_z \geqslant -k. \tag{40}$$

El caso de igualdad se tendrá si y sólo si $\eta|m'_z\rangle = 0$. Análogamente, de (35)

$$\langle m'_z|\eta\bar{\eta}|m'_z\rangle = (\beta - m'^2_z - \hbar m'_z)\langle m'_z|m'_z\rangle,$$

lo que demuestra que

$$\beta - m'^2_z - \hbar m'_z \geqslant 0$$

o bien

$$k \geqslant m'_z \geqslant -k-\hbar,$$

verificándose una igualdad si y sólo si $\bar{\eta}|m'_z\rangle = 0$. Esta condición junto con (40) nos muestra que $k \geqslant 0$ y que

$$k \geqslant m'_z \geqslant -k, \tag{41}$$

siendo $m'_z = k$ si $\bar{\eta}|m'_z\rangle = 0$, y $m'_z = -k$ si $\eta|m'_z\rangle = 0$.

De (37)

$$m_z \eta\, |m'_z\rangle = (\eta m_z - \hbar\eta)\, |m'_z\rangle = (m'_z - \hbar)\bar{\eta}\, |m'_z\rangle.$$

Si $m'_z \neq -k$, $\eta|m'_z\rangle$ es distinto de cero y por tanto es un autoket de m_z perteneciente al autovalor $m'_z-\hbar$. Análogamente, si $m'_z-\hbar \neq -k, m'_z-2\hbar$ también es autovalor de m_z, y así sucesivamente. Así obtenemos la sucesión de autovalores m'_z, $m'_z-\hbar$, $m'_z-2\hbar$, ... que en virtud de (41) ha de acabar forzosamente, y sólo puede acabar con el valor $-k$. A su vez, la ecuación compleja conjugada de (37)

$$m_z \bar{\eta}\, |m'_z\rangle = (\bar{\eta} m_z + \hbar\bar{\eta})\, |m'_z\rangle = (m'_z + \hbar)\bar{\eta}\, |m'_z\rangle,$$

demuestra que $m'_z + \hbar$ es también autovalor de m_z salvo si $\bar{\eta}|m'_z\rangle = 0$, en cuyo caso $m'_z = k$. Reiterando el procedimiento obtenemos otra sucesión de autovalores m'_z, $m'_z + \hbar$, $m'_z + 2\hbar$, ..., que en virtud de (41) tiene que terminar forzosamente, y sólo puede terminar con el valor k. Podemos concluir que $2k$ tiene que ser un múltiplo entero de $\hbar$ y que los autovalores de m_z son

$$k,\ k-\hbar,\ k-2\hbar,\ \ldots, -k+\hbar,\ -k. \tag{42}$$

Por simetría, los autovalores de m_x y m_y han de ser los mismos. Todos estos autovalores son múltiplos enteros o bien múltiplos impares semienteros de $\hbar$, según que $2k$ sea un mútiplo par o impar de $\hbar$.

Sea $|\text{max}\rangle$ un autoket de m_z perteneciente al autovalor máximo k, con lo que

$$\bar{\eta}|\text{max}\rangle = 0, \tag{43}$$

y formemos la sucesión de kets

$$|\text{max}\rangle,\quad \eta|\text{max}\rangle,\quad \bar{\eta}^2|\text{max}\rangle,\quad \ldots,\quad \bar{\eta}^{2k/\hbar}|\text{max}\rangle. \tag{44}$$

Todos estos kets son autokets de m_z pertenecientes respectivamente a la sucesión de autovalores (42). El conjunto de kets (44) es tal que al aplicar el operador $\bar{\eta}$ a cualquiera de ellos se obtiene un ket dependiente del conjunto (η aplicado al último da cero) y según (36) y (43) vemos que al

aplicar $\bar{\eta}$ a uno cualquiera de ellos obtenemos también un ket que depende del conjunto. Todas las variables dinámicas del sistema que estamos considerando se pueden expresar en función de η y de $\bar{\eta}$, y por tanto, el conjunto de kets (44) es un conjunto completo. Hay un solo ket perteneciente a cada autovalor (42) de m_z, por lo tanto m_z constituye por sí solo un conjunto completo de observables.

Es conveniente definir el módulo del vector momento angular $\mathbf{m}$ igual a k definido por (39) y no igual a $\beta^{\frac{1}{2}}$, ya que los valores posibles de k son

$$0,\ \tfrac{1}{2}\hbar,\ \hbar,\ \tfrac{3}{2}\hbar,\ 2\hbar,\ \ldots, \tag{45}$$

hasta infinito, mientras que los de $\beta^{\frac{1}{2}}$ constituyen un conjunto de números más complicado.

Para un sistema dinámico que tenga otras variables dinámicas además de m_x, m_y y m_z pueden existir variables que no conmuten con β. En ese caso β ya no será un número sino un operador lineal. Para todo momento angular orbital (22) ocurrirá esto, pues x, y, z, p_x, p_y y p_z no conmutan con β. Supondremos que β es siempre un observable, y que por tanto, se puede definir (39), utilizando la función raíz cuadrada positiva, que también será un observable. Diremos que el k así definido es la magnitud del vector momento angular en el caso general. El análisis anterior que empleamos para obtener los autovalores de m_z sigue siendo válido en este caso si sustituimos $|m'_z\rangle$ por un autoket $|k'm'_z\rangle$ común a los observables que conmutan k y m_z, resultando que los autovalores posibles de k son los números (45), y que para cada autovalor k' de k los autovalores de m_z son los números (42) sustituyendo k por k'. He aquí un ejemplo de un fenómeno con el que no nos habíamos encontrado anteriormente, a saber, que para dos observables que conmutan, los autovalores del uno dependen de qué autovalor del otro consideremos. Este fenómeno se puede interpretar como que ambos observables no son completamente independientes, sino que parcialmente son función uno de otro. El número de autokets independientes comunes a k y m_z que pertenecen a los autovalores k' y m'_z no ha de depender de m'_z, pues para cada $|k'm'_z\rangle$ independiente podemos obtener un $|k'm''_z\rangle$ también independiente con cualquier valor de m''_z de la sucesión (42), sin más que multiplicar $|k'm'_z\rangle$ por una potencia adecuada de η o de $\bar{\eta}$.

Como ejemplo consideremos un sistema dinámico que posea dos momentos angulares $\mathbf{m}_1$ y $\mathbf{m}_2$, que conmutan entre sí. Si no existe ninguna otra variable dinámica, todas las variables dinámicas conmutarán con las cantidades k_1 y k_2 de $\mathbf{m}_1$ y $\mathbf{m}_2$, luego, tanto k_1 como k_2 son números. Sin embargo, la magnitud k del momento angular resultante $\mathbf{M} = \mathbf{m}_1 + \mathbf{m}_2$ no es un número (ya que no conmuta con las componentes de $\mathbf{m}_1$ y de $\mathbf{m}_2$) y resulta interesante calcular los autovalores de K. El modo más sencillo de hacerlo se basa en un procedimiento de contar kets independientes.

Hay un solo autoket independiente común a m_{1z} y m_{2z} que pertenezca a un autovalor m'_{1z} para cada uno de los valores k_1, $k_1 - \hbar$, $k_1 - 2$, ..., $-k_1$ y a un autovalor m'_{2z} para cada uno de los valores k_2, $k - \hbar$, $k_2 - 2\hbar$, ..., $-k_2$, y dicho ket es también autoket de M_z perteneciente al autovalor $M'_z = m'_{1z} + m'_{2z}$. Los valores posibles de M'_z serán, por tanto, $k_1 + k_2$, $k_1 + k_2 - \hbar$, $k_1 + k_2 - 2\hbar$, ..., $-k_1 - k_2$, y el número de veces que ocurrirá cada uno de ellos viene dado por el siguiente esquema (para concretar supondremos $k_1 \geqslant k_2$),

$$\left.\begin{array}{l} k_1+k_2,\ k_1+k_2-\hbar,\ k_1+k_2-2\hbar,\ldots,k_1-k_2,k_1-k_2-\hbar,\ldots \\ \quad 1 \qquad\quad 2 \qquad\qquad 3 \qquad\ldots\ 2k_2+1 \quad 2k_2+1 \quad \ldots \\ \qquad\qquad\qquad\qquad \ldots\ -k_1+k_2,-k_1+k_2-\hbar,\ldots,-k_1-k_2 \\ \qquad\qquad\qquad\qquad \ldots\ 2k_2+1 \qquad 2k_2 \quad \ldots \quad 1 \end{array}\right\} \quad (46)$$

Ahora bien, para cada autovalor K' de K, M_z tendrá los autovalores K', $K' - \hbar$, $K' - 2\hbar$, ..., $-K'$ y para cada par de autovalores de K y M_z existirá el mismo número de autokets independientes comunes a K y a M_z. El número total de autokets de M_z independientes que pertenecen a un autovalor M'_z tiene que ser el mismo tanto si los elegimos entre los autokets comunes a m_{1z} y m_{2z} como si los elegimos entre los autokets comunes a K y M_z, luego, en ambos casos viene dado por el esquema (46). De aquí se deduce que los autovalores de K son

$$k_1 + k_2, \quad k_1 + k_2 - \hbar, \quad k_1 + k_2 - 2\hbar, \quad \ldots, \quad k_1 - k_2, \tag{47}$$

y que para cada uno de estos autovalores de K y un autovalor de M_z compatible con él existe un único autoket independiente común a K y a M_z.

Hemos de considerar la acción de las rotaciones sobre los autokets de las variables de momento angular. Sea $|M'_z\rangle$ un autoket cualquiera de la componente z del momento angular total de un sistema dinámico general, y apliquémosle un pequeño giro de ángulo $\delta\phi$ en torno al eje z. Según (32) se transformará en

$$(1 + \delta\phi r_z)\,|M'_z\rangle = (1 - i\delta\phi M_z/\hbar)\,|M'_z\rangle$$

Esto equivale a

$$(1 - i\delta\phi M'_z\hbar/)\,|M'_z\rangle = e^{-i\delta\phi M'_z/\hbar}\,|M'_z\rangle$$

para el primer orden en $\delta\phi$. Luego $|M'_z\rangle$ queda multiplicado por el factor numérico $e^{-i\delta\phi M'_z/\hbar}$. Aplicando sucesivamente pequeños giros resulta que una rotación de ángulo ϕ alrededor del eje z transforma a $|M'_z\rangle$ en sí mismo multiplicado por $e^{-i\phi M'_z/\hbar}$. Tomando $\phi = 2\pi$ resulta que un giro de 2π alrededor del eje z deja invariante a $|M'_z\rangle$ cuando el autovalor M'_z es múltiplo entero de $\hbar$, y le cambia el signo si M'_z es múltiplo impar semientero de $\hbar$. Ahora consideremos un autoket $|K'\rangle$ de la magnitud del momento angular total K. Si el autovalor K' es un múltiplo entero de $\hbar$, los

autovalores posibles de M_z serán todos múltiplos enteros de $\hbar$, y un giro de 2π alrededor del eje z dejará invariante a $|K'\rangle$. Por el contrario, si K' es un múltiplo impar semientero, los autovalores posibles de M_z serán todos múltiplos impares semienteros de $\hbar$, y un giro de 2π en torno al eje z cambiará el signo de $|K'\rangle$. Por simetría, un giro de 2π en torno a cualquier eje tendrá los mismos efectos sobre $|K'\rangle$ que un giro de 2π en torno al eje z. Así llegamos a la conclusión general de que *un giro de 2π alrededor de cualquier eje deja invariante a un ket o bien le cambia el signo, según que pertenezca a autovalores de la magnitud del momento angular total que sean múltiplos enteros o múltiplos impares semienteros de $\hbar$.* Desde luego, el estado no cambia con el giro de 2π, pues un estado sigue siendo el mismo aunque cambiemos el signo del ket que lo representa.

Un ket de un sistema dinámico que sólo posea momentos angulares orbitales, tiene que ser invariante frente a un giro de 2π en torno de un eje cualquiera, pues para dicho sistema podemos considerar la representación de Schrödinger en la que las coordenadas de todas las partículas son diagonales, y en ella el representante de un ket se transforma en sí mismo en dicho giro. En consecuencia, *los autovalores de la magnitud de un momento angular orbital son siempre múltiplos enteros de $\hbar$.* Asimismo los autovalores de las componentes de un momento angular orbital serán también múltiplos enteros de $\hbar$ siempre. Para los momentos angulares de spin no existe una representación de Schrödinger y, por tanto, pueden existir ambos tipos de autovalores.

37. *El spin del electrón*

Los electrones así como también otras partículas fundamentales (protones, neutrones) tienen un spin cuyo módulo vale $\frac{1}{2}\hbar$. Este valor se encontró experimentalmente, pero además existen razones teóricas que nos indican que tal valor del spin es el más elemental, incluso más que el valor cero (véase capítulo XI). Así pues, el estudio de este spin particular es de gran importancia.

Para tratar con un momento angular cuyo módulo vale $\frac{1}{2}\hbar$, es conveniente poner

$$\mathbf{m} = \hbar\tfrac{1}{2}\boldsymbol{\sigma}. \tag{48}$$

Según (27), las componentes del vector $\boldsymbol{\sigma}$ verifican

$$\left.\begin{aligned} \sigma_y\,\sigma_z - \sigma_z\,\sigma_y &= 2i\sigma_x, \\ \sigma_z\,\sigma_x - \sigma_x\,\sigma_z &= 2i\sigma_y, \\ \sigma_x\,\sigma_y - \sigma_y\,\sigma_x &= 2i\sigma_z. \end{aligned}\right\} \tag{49}$$

Los autovalores de m_z son $\frac{1}{2}\hbar$ y $-\frac{1}{2}\hbar$, luego los de σ_z serán 1 y -1, y

σ_z^2 no tendrá más que un único autovalor que será el 1. Por tanto, σ_z^2 ha de ser igual a 1 y análogamente ha de ocurrir con σ_x^2 y σ_y^2,

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1. \tag{50}$$

Podemos simplificar las ecuaciones (49) y (50) utilizando artificios del álgebra no conmutativa a la que obedecen. Según (50) tenemos

$$\sigma_y^2 \sigma_z - \sigma_z \sigma_y^2 = 0$$

o bien

$$\sigma_y(\sigma_y \sigma_z - \sigma_z \sigma_y) + (\sigma_y \sigma_z - \sigma_z \sigma_y)\sigma_y = 0$$

que, con la ayuda de la primera de las ecuaciones (49), se convierte en

$$\sigma_y \sigma_x + \sigma_x \sigma_y = 0.$$

Esto quiere decir que $\sigma_x \sigma_y = - \sigma_y \sigma_x$. Dos variables dinámicas o dos operadores lineales como éstos que salvo el signo negativo verifican la propiedad conmutativa de la multiplicación se dice que *anticonmutan*. Así decimos que σ_x anticonmuta con σ_y. Por simetría cada una de las tres variables dinámicas σ_x, σ_y y σ_z anticonmutará con las demás. Las ecuaciones (49) se pueden escribir

$$\left.\begin{aligned} \sigma_y \sigma_z &= i\sigma_x = - \sigma_z \sigma_y, \\ \sigma_z \sigma_x &= i\sigma_y = - \sigma_x \sigma_z, \\ \sigma_x \sigma_y &= i\sigma_z = - \sigma_y \sigma_x, \end{aligned}\right\} \tag{51}$$

y asimismo de (50)

$$\sigma_x \sigma_y \sigma_z = i. \tag{52}$$

Las ecuaciones (50), (51) y (52) son las ecuaciones fundamentales para las variables de spin σ que representan un spin cuyo módulo es $\frac{1}{2}\hbar$.

Establezcamos una representación matricial para las variables σ en la que σ_z sea diagonal. Si no existe ninguna otra variable dinámica independiente en nuestro sistema dinámico más que las m o las σ, σ_z constituye por sí sola un conjunto completo de observables, pues de la forma de las ecuaciones (50) y (51) inferimos la imposibilidad de construir ninguna variable dinámica nueva a partir de σ_x, σ_y y σ_z que conmute con σ_z. Como los elementos de matriz diagonales de la representación matricial de σ_z han de ser sus autovalores 1 y -1, la matriz será

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Sea

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$$

la matriz que representa a σ_x. Dicha matriz ha de ser hermítica y, por

tanto, a_1 y a_4 han de ser reales, y a_2 y a_3 han de ser números complejos conjugados. La ecuación $\sigma_z \sigma_x = -\sigma_x \sigma_z$ nos da

$$\begin{pmatrix} a_1 & a_2 \\ -a_3 & -a_4 \end{pmatrix} = -\begin{pmatrix} a_1 & -a_2 \\ a_3 & -a_4 \end{pmatrix}$$

luego $a_1 = a_4 = 0$. Por lo tanto, la representación matricial de σ_x será de forma

$$\begin{pmatrix} 0 & a_2 \\ a_3 & 0 \end{pmatrix}$$

La ecuación $\sigma_x^2 = 1$ nos dice que $a_2 a_3 = 1$. Por tanto, como a_2 y a_3 son números complejos conjugados, han de ser de la forma $e^{i\alpha}$ y $e^{-i\alpha}$ respectivamente, siendo α un número real. Así pues, σ_x estará representada por una matriz de la forma

$$\begin{pmatrix} 0 & e^{i\alpha} \\ e^{-i\alpha} & 0 \end{pmatrix}$$

Análogamente veríamos que σ_y también ha de estar representada por una matriz de esta forma. Eligiendo convenientemente los factores de fase que no están unívocamente determinados por la simple condición de que σ_z sea diagonal, podemos hacer que σ_x esté representada por la matriz

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

El representante de σ_y viene así determinado por la ecuación $\sigma_y = i\sigma_x \sigma_z$. Obtenemos finalmente las tres matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{53}$$

que representan respectivamente a σ_x, σ_y y σ_z, y que verifican todas las relaciones algebraicas dadas por (49), (50), (51) y (52). La componente del vector $\boldsymbol{\sigma}$ en una dirección cualquiera dada por los cosenos directores l, m, n, o sea $l\sigma_x + m\sigma_y + n\sigma_z$ estará representada por

$$\begin{pmatrix} n & l - im \\ l + im & -n \end{pmatrix} \tag{54}$$

El representante de un ket constará de dos números, que corresponderán a los dos valores $+1$ y -1 de σ'_z. Dicho par de números constituyen una función $f_\beta(\sigma'_z)$ cuyo dominio de definición está constituido únicamente del par de puntos $+1$ y -1. El estado en que σ_z tiene el valor uno estará representado por la función $f_\alpha(\sigma'_z)$ cuyos valores serán 1, 0, y el estado en que σ_z tiene el valor -1 vendrá representado por la función

$f_\beta(\sigma'_z)$ cuyos valores serán 0, 1. Cualquier par de números se puede expresar como combinación lineal de los dos anteriores. Así pues, *todo estado se puede obtener por superposición de los dos estados para los que* σ_z *toma respectivamente los valores* $+1$ y -1. Por ejemplo, el estado en el cual la componente de σ_z en la dirección l, m, n, representada por la matriz (54), tiene el valor $+1$ estará caracterizado por el par de números a y b que verifiquen.

$$\begin{pmatrix} n & l-im \\ l+im & -n \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

o bien

$$na + (l-im)b = a,$$
$$(l+im)a - nb = b.$$

Por tanto,

$$\frac{a}{b} = \frac{l-im}{1-n} = \frac{1+n}{l+im}.$$

Este estado puede ser considerado como la superposición de los dos estados en los cuales σ_z es igual a $+1$ y a -1, cuyos pesos relativos en la superposición serán

$$|a|^2 : |b|^2 = |l-im|^2 : (1-n)^2 = 1+n : 1-n. \tag{55}$$

Para la descripción completa de un electrón (u otra partícula elemental de spin $\frac{1}{2}\hbar$) necesitamos las variables dinámicas de spin σ, relacionadas con el momento angular de spin a través de (48), junto con las coordenadas cartesianas x, y, z y los momentos p_x, p_y, p_z. Las variables dinámicas de spin conmutan con dichas coordenadas y momentos. Por tanto, un conjunto completo de observables que conmutan para un sistema constituido por un único electrón puede ser x, y, z y σ_z. En una representación en la que todos ellos sean diagonales, el representante de cualquier estado será una función de las cuatro variables x', y', z' σ'_z Como σ'_z no puede tomar más que los dos valores 1 y -1, la función de cuatro variables es equivalente a dos funciones de tres variables.

$$\langle x'y'z'|\rangle_+ = \langle x', y', z', +1|\rangle, \qquad \langle x'y'z'|\rangle_- = \langle x', y', z', -1|\rangle. \tag{56}$$

Así pues, *la presencia del spin se puede considerar o bien introduciendo una nueva variable en el representante de un estado o bien dando dos componentes al representante.*

38. *Movimiento en un campo de fuerzas central*

Un átomo está constituido por un núcleo de gran masa, cargado positivamente, y un conjunto de electrones que se mueven alrededor de él bajo la influencia de las fuerzas de atracción del núcleo y las repulsiones mutuas entre ellos. El estudio exacto de este sistema dinámico es un problema matemático muy difícil. Sin embargo, se puede obtener alguna luz sobre el comportamiento general de todo el sistema haciendo una aproximación grosera que consiste en considerar que cada electrón se mueve independientemente en un cierto campo de fuerzas *central* al que se superpone un cierto valor medio de las fuerzas ejercidas por los demás electrones. Así pues, el problema del movimiento de una partícula en un campo de fuerzas central, que vamos a estudiar ahora, constituye un elemento fundamental en la teoría del átomo.

Sean x, y, z las coordenadas cartesianas de la partícula referidas a un sistema de ejes cuyo origen es el centro de fuerzas, y p_x, p_y, p_z las componentes del momento correspondientes. Despreciando la aproximación relativista, el hamiltoniano será de la forma

$$H = 1/2m \cdot (p_x^2 + p_y^2 + p_z^2) + V, \tag{57}$$

siendo la energía potencial V función únicamente de $(x^2 + y^2 + z^2)$. Para desarrollar la teoría es conveniente introducir variables dinámicas polares. En primer lugar, introducimos el radio r mediante la raíz cuadrada positiva

$$r = (x^2 + y^2 + z^2)^{\frac{1}{2}}.$$

Sus autovalores van desde 0 hasta ∞. Si calculamos sus P. B. con p_x, p_y, y p_z mediante la fórmula (32) de § 22 obtenemos

$$[r, p_x] = \frac{\partial r}{\partial x} = \frac{x}{r}, \qquad [r, p_y] = \frac{y}{r}, \qquad [r, p_z] = \frac{z}{r},$$

que son los mismos que se obtienen en la teoría clásica. Asimismo, introducimos también la variable dinámica p_r mediante

$$p_r = r^{-1}(xp_x + yp_y + zp_z). \tag{58}$$

Su P. B. con r viene dado por

$$\begin{aligned} r[r, p_r] &= [r, rp_r] = [r, xp_x + yp_y + zp_z] \\ &= x[r, p_x] + y[r, p_y] + z[r, p_z] \\ &= x \cdot x/r + y \cdot y/r + z \cdot z/r = r. \end{aligned}$$

Luego,

$$[r, p_r] = 1$$

o bien,

$$rp_r - p_r r = i\hbar.$$

La relación de conmutación entre r y p_r es precisamente la que corresponde a un par de variables canónicas conjugadas, es decir, la ecuación (10) de § 22. Luego, p_r hace las veces de momento canónico conjugado de la coordenada r. Pero no es exactamente igual a él, pues no es real; su complejo conjugado es

$$\begin{aligned}\bar{p}_r = (p_x x + p_y y + p_z z)r^{-1} &= (xp_x + yp_y + zp_z - 3i\hbar)r^{-1} \\ &= (rp_r - 3i\hbar)r^{-1} = p_r - 2i\hbar r^{-1}. \qquad (59)\end{aligned}$$

Por tanto, $p_r - i\hbar r^{-1}$ es real y es el verdadero momento canónico conjugado de r.

El momento angular $\mathbf{m}$ de la partícula alrededor del origen viene dado por (22) y su magnitud k por (39). Puesto que r y p_r son escalares, conmutan con $\mathbf{m}$ y por tanto también con k. Podemos expresar el hamiltoniano en función de r, p_r y k. Si $\sum_{xyz}$ significa suma respecto a las permutaciones cíclicas de los subíndices x, y, z, tenemos

$$\begin{aligned}k(k+\hbar) &= \sum_{xyz} m_z^2 = \sum_{xyz} (xp_y - yp_x)^2 \\ &= \sum_{xyz} (xp_y xp_y + yp_x yp_x - xp_y yp_x - yp_x xp_y) \\ &= \sum_{xyz} (x^2p_y^2 + y^2p_x^2 - xp_x p_y y - yp_y p_x x + x^2p_x^2 - xp_x p_x x - 2i\hbar xp_x) \\ &= (x^2+y^2+z^2)(p_x^2+p_y^2+p_z^2) - \\ &\qquad -(xp_x+yp_y+zp_z)(p_x x+p_y y+p_z z+2i\hbar) \\ &= r^2(p_x^2+p_y^2+p_z^2) - rp_r(\bar{p}_r r + 2i\hbar) \\ &= r^2(p_x^2+p_y^2+p_z^2) - rp_r^2 r.\end{aligned}$$

según (59). Luego

$$H = \frac{1}{2m}\left(\frac{1}{r}p_r^2 r + \frac{k(k+\hbar)}{r^2}\right) + V. \qquad (60)$$

Esta forma de H es tal que k conmuta no sólo con H como debiera ocurrir por ser k una constante del movimiento, sino también con todas las variables dinámicas que aparecen en H, a saber, r, p_r y V, siendo esta última, función de r. Por tanto, el problema se resuelve fácilmente, pues podemos considerar un autoestado de k perteneciente al autovalor k' y sustituir k por k' en (60) resultando con ello un problema con un solo grado de libertad r.

Introduzcamos la representación de Schrödinger con x, y, z diagonales. Entonces p_x, p_y y p_z son equivalentes a los operadores $-i\hbar\partial/\partial x$, $-i\hbar\partial/\partial y$ y $-i\hbar\partial/\partial z$ respectivamente. Los estados vendrán representados por una

función de onda $\psi(xyzt)$ que verifique la ecuación de ondas de Schrödinger (7) de § 27, que en este caso con H dado por (57) será

$$i\hbar\frac{\partial\psi}{\partial t} = \left\{-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial^2 x}+\frac{\partial^2}{\partial^2 y}+\frac{\partial^2}{\partial^2 z}\right)+V\right\}\psi. \tag{61}$$

Podemos pasar de las coordenadas cartesianas x, y, z a las coordenadas polares r, θ, ϕ, mediante las ecuaciones

$$\left.\begin{aligned} x &= r \operatorname{sen}\theta\cos\phi,\\ y &= r\operatorname{sen}\theta\operatorname{sen}\phi,\\ z &= r\cos\theta,\end{aligned}\right\} \tag{62}$$

y expresar la función de ondas en función de las coordenadas polares en la forma $\psi(r\theta\phi t)$. De las ecuaciones (62) resulta la siguiente ecuación entre operadores

$$\frac{\partial}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial}{\partial x}+\frac{\partial y}{\partial r}\frac{\partial}{\partial y}+\frac{\partial z}{\partial r}\frac{\partial}{\partial z} = \frac{x}{r}\frac{\partial}{\partial x}+\frac{y}{r}\frac{\partial}{\partial y}+\frac{z}{r}\frac{\partial}{\partial z},$$

que comparada con (58) nos dice que $p_r = -i\hbar\partial/\partial r$. Así, pues, la ecuación de ondas de Schrödinger con el hamiltoniano en la forma (60) es la siguiente

$$i\hbar\frac{\partial\psi}{\partial t} = \left\{\frac{\hbar^2}{2m}\left(-\frac{1}{r}\frac{\partial^2}{\partial r^2}r+\frac{k(k+\hbar)}{\hbar^2 r^2}\right)+V\right\}\psi. \tag{63}$$

Aquí k es un cierto operador lineal que como conmuta con r y $\partial/\partial r$ sólo puede depender de θ, ϕ, $\partial/\partial\theta$ y $\partial/\partial\phi$. A partir de la fórmula

$$k(k+\hbar) = m_x^2+m_y^2+m_z^2, \tag{64}$$

que se deduce de (39), y de la fórmula (62), podemos calcular la expresión $k(k+\hbar)$ que resulta ser

$$\frac{k(k+\hbar)}{\hbar^2} = -\frac{1}{\operatorname{sen}\theta}\frac{\partial}{\partial\theta}\operatorname{sen}\theta\frac{\partial}{\partial\theta}-\frac{1}{\operatorname{sen}^2\theta}\frac{\partial^2}{\partial\phi^2} \tag{65}$$

Este operador es muy conocido en física matemática. Sus autofunciones son los llamados *armónicos esféricos*, y sus autovalores son $n(n+1)$ siendo n un número entero. Así pues, la teoría de los armónicos esféricos constituye otra prueba de que los autovalores de k son múltiplos enteros de $\hbar$.

La función de ondas para un autoestado de k perteneciente al autovalor $n\hbar$ (n es un entero no negativo) será de la forma

$$\psi = r^{-1}\chi(rt)S_n(\theta\phi), \tag{66}$$

donde $S_n(\theta\phi)$ verificará la ecuación

$$k(k+\hbar)S_n(\theta\phi) = n(n+1)\hbar^2 S_n(\theta\phi), \tag{67}$$

y según (65) será un armónico esférico de orden n. El factor r^{-1} de (66) lo hemos puesto por conveniencia. Sustituyendo (66) en (63) resulta la siguiente ecuación para χ

$$i\hbar\frac{\partial\chi}{\partial t} = \left\{\frac{\hbar^2}{2m}\left(-\frac{\partial^2}{\partial r^2}+\frac{n(n+1)}{r^2}\right)+V\right\}\chi. \tag{68}$$

Si el estado es un estado estacionario perteneciente al valor de la energía H', χ será de la forma

$$\chi(rt) = \chi_0(r)e^{-iH't/\hbar}$$

y (68) se reducirá a

$$H'\chi_0 = \left\{\frac{\hbar^2}{2m}\left(-\frac{d^2}{dr^2}+\frac{n(n+1)}{r^2}\right)+V\right\}\chi_0. \tag{69}$$

Esta ecuación se puede utilizar para determinar los niveles de energía H' del sistema. Para cada solución χ_0 de (69) correspondiente a un determinado n, existirán $2n+1$ estados independientes, pues existen $2n+1$ valores distintos que puede tomar una de las componentes del momento angular, por ejemplo m_z.

La probabilidad de que la partícula esté en un elemento de volumen $dxdydz$ es proporcional a $|\psi|^2dxdydz$. Poniendo ψ en la forma (66) esto equivale a $r^{-2}|\chi|^2|S_n|^2dxdydz$. La probabilidad de que la partícula esté comprendida entre las dos esferas de radios r y $r+dr$ será proporcional a $|\chi|^2dr$. Así resulta evidente que al resolver la ecuación (68) o la (69) hemos de imponer una condición de contorno a la función χ para $r=0$, y que la integral $\int_0|\chi|^2dr$ converja en torno al origen. Si la integral en cuestión no fuera convergente, la función de ondas representaría un estado para el que habría una probabilidad infinita de que las partículas estuvieran en el origen, lo que es inadmisible desde el punto de vista físico.

Sin embargo, la condición obtenida para $r=0$ mediante la consideración del significado probabilístico no es suficientemente restrictiva. Al exigir que la función de onda que se obtiene resolviendo el problema en coordenadas polares (63) sea realmente solución de la ecuación de ondas en coordenadas cartesianas (61), obtenemos una condición más restrictiva. Consideremos el caso $V=0$, que constituye el problema de la partícula libre. La ecuación (61) aplicada a un estado estacionario de energía $H'=0$ es ahora

$$\nabla^2\psi = 0, \tag{70}$$

donde ∇^2 representa el operador laplaciano $\partial^2/\partial x^2+\partial^2/\partial y^2+\partial^2/\partial z^2$, y la (63) es

$$\left(\frac{1}{r}\frac{\partial^2}{\partial r^2}r-\frac{k(k+\hbar)}{\hbar^2r^2}\right)\psi = 0. \tag{71}$$

para $k = 0$, $\psi = r^{-1}$ es una solución particular de (71). Sin embargo, esta función no es solución de (70), pues aunque $\nabla^2 r^{-1}$ se anula para todo valor finito de r, su integral extendida a un volumen que contenga el origen vale -4π (como puede verse transformando dicha integral de volumen en una integral de superficie mediante el teorema de Gauss), y, por tanto,

$$\nabla^2 r^{-1} = -4\pi\,\delta(x)\delta(y)\delta(z). \tag{72}$$

Luego, no toda solución de (71) es solución de (70), y más en general, no toda solución de (63) es solución de (61). Para que al sustituir una solución de (63), no aparezca una función δ en el segundo miembro de (61), tal como ocurre en el segundo miembro de (72), hemos de imponer la condición de que las soluciones de (63) no tiendan a infinito con la misma rapidez que r^{-1} cuando $r \to 0$. La ecuación (63) sólo es equivalente a la (61) cuando se le añade esta condición suplementaria. Así obtenemos la condición de contorno $r\psi \to 0$ o $\chi \to 0$ cuando $r \to 0$.

También hemos de imponer condiciones de contorno a la función de ondas para $r = \infty$. Si solamente nos interesan los estados 'ligados', es decir, los estados en que la partícula no pueda alcanzar el infinito, hemos de limitarnos a considerar el caso en que la integral $\int |\chi(r)|^2 dr$ sea convergente en el infinito. Sin embargo, los estados ligados no son los únicos estados que pueden existir desde el punto de vista físico, pues también podemos considerar casos en los que la partícula venga desde el infinito, sea dispersada (scattered) por el campo de fuerzas central, y se dirija después nuevamente al infinito. Para tales estados la función de ondas puede ser finita para $r \to \infty$. Estos estados serán estudiados en el capítulo VIII cuando consideremos los problemas de colisión. Pero, en cualquier caso, la función de onda no debe tender a infinito para $r \to \infty$, pues representaría un estado sin ningún significado físico.

39. *Niveles de energía del átomo de hidrógeno*

El análisis anterior se puede aplicar al problema del átomo de hidrógeno despreciando la mecánica relativista y el spin del electrón. La energía potención V en este caso vale † $-e^2/r$, y la ecuación (69) se convierte en

$$\left\{ \frac{d^2}{dr^2} - \frac{n(n+1)}{r^2} + \frac{2me^2}{\hbar^2}\frac{1}{r} \right\}\chi_0 = -\frac{2mH'}{\hbar^2}\chi_0. \tag{73}$$

El estudio completo de esta ecuación fue hecho por Schrödinger.‡ Aquí obtendremos sus autovalores H' por un razonamiento elemental.

† Aquí e representa la carga del electrón con signo contrario, y no debe confundirse con el número e utilizando como base en las exponenciales.

‡ SCHRÖDINGER, *Ann. d. Physik*, 79 (1926), 361.

Es conveniente poner

$$\chi_0 = f(r)e^{-r/a}, \tag{74}$$

e introducir la nueva función $f(r)$, siendo el coeficiente a igual a una de las dos raíces cuadradas

$$a = \pm \sqrt{(-\hbar^2/2mH')}. \tag{75}$$

La ecuación (73) escrita en función de $f(r)$ es

$$\left\{ \frac{d^2}{dr^2} - \frac{2}{a}\frac{d}{dr} - \frac{n(n+1)}{r^2} + \frac{2me^2}{\hbar^2}\frac{1}{r} \right\} f(r) = 0. \tag{76}$$

Busquemos una solución de la ecuación de la forma de una serie de potencias

$$f(r) = \sum_s c_s r^s, \tag{77}$$

en la que los valores de s consecutivos difieran en una unidad, aunque no exigimos que hayan de ser enteros. Sustituyendo (77) en (76) resulta

$$\sum_s c_s\{s(s-1)r^{s-2} - (2s/a)r^{s-1} - n(n+1)r^{s-2} + (2me^2/\hbar^2)r^{s-1}\} = 0,$$

que al igualar a cero el coeficiente de r^{s-2} da la siguiente relación entre los coeficientes c_s consecutivos

$$c_s[s(s-1) - n(n+1)] = c_{s-1}[2(s-1)/a - 2me^2/\hbar^2]. \tag{78}$$

En la sección anterior vimos que únicamente debíamos considerar autofunciones χ que tendieran a cero con r y, por tanto, según (74), $f(r)$ tiene que tender a cero con r. Luego, la serie (77) tiene que terminarse por el lado de las s pequeñas, y el mínimo valor de s ha de ser mayor que cero. Pero los únicos valores mínimos de s posibles son los que anulan el coeficiente de c_s en (78), es decir, $n+1$ y $-n$, el segundo de los cuales es negativo o nulo. En consecuencia, el mínimo valor de s ha de ser $n+1$. Puesto que n es siempre entero, los valores de s habrán de ser también enteros. En general, la serie (77) no tiene por qué terminarse para valores grandes de s. Para grandes valores de s, la relación entre términos consecutivos será, según (78),

$$\frac{c_s}{c_{s-1}} r = \frac{2r}{sa}$$

Por tanto, la serie (77) convergerá siempre, pues la relación entre términos consecutivos para valores grandes de s es la misma que la de la serie

$$\sum_s \frac{1}{s!}\left(\frac{2r}{a}\right)^s \tag{79}$$

que converge hacia $e^{2r/a}$.

Ahora hemos de investigar el comportamiento de la solución χ_0 para valores grandes de r. Hemos de distinguir entre valores de H' positivos y valores de H' negativos, el número a de (75) será real. Tomemos, por ejemplo, el valor positivo para a. Entonces para $r \to \infty$ la suma de la serie (77) tenderá a infinito según la misma ley que la serie (79), o sea, como $e^{2r/a}$. Por tanto, según (74), χ_0 tendería a infinito según la ley $e^{r/a}$, no representando ningún estado con sentido físico. Luego, para H' negativo no hay, en general, solución de (73). Sin embargo, existe una excepción y es cuando la serie (77) se termina para valores grandes de s, en cuyo caso se verifican todas las ecuaciones de contorno. La condición para que ocurra eso es que se anule el coeficiente de c_{s-1} en (78) para algún valor del subíndice $s - 1$ que no sea menor que su valor $n + 1$, lo que equivale a la condición

$$\frac{s}{a} - \frac{me^2}{\hbar^2} = 0$$

para algún valor de s no inferior a $n + 1$. Con ayuda de (75) esta condición se transforma en

$$H' = -\frac{me^4}{2s^2\hbar^2}, \tag{80}$$

que es una condición para los niveles de energía H'. Puesto que s puede ser cualquier número entero positivo, la fórmula (80) da un conjunto discreto de posibles niveles de energía negativa en el átomo de hidrógeno. Dichos niveles están de acuerdo con la experiencia. Para cada nivel (excepto para el más bajo $s = 1$) existen varios estados independientes correspondientes a los diversos valores posibles de n, que pueden ser igual a cualquier número entero positivo o nulo menor que s. Esta multiplicidad de estados pertenecientes a un mismo valor de la energía hay que añadirla a la estudiada en la sección anterior, debida a los diversos valores posibles de una componente del momento angular que existe en cualquier campo de fuerzas central. La multiplicidad debida a n tiene lugar únicamente con una ley de fuerza del inverso del cuadrado de la distancia, e incluso en este caso desaparece al tener en cuenta la mecánica relativista, como veremos en el capítulo XI. La solución χ_0 de (73) para valores H' de (80) tiende a cero exponencialmente para $r \to \infty$ y, por tanto, representa un estado ligado (que corresponde a las órbitas elípticas de la teoría de Bohr).

Para valores de H' positivos, el número a de (75) será imaginario puro. En tal caso, la suma de la serie (77), que para valores grandes de r es análoga a la serie (79), conservará un valor finito cuando $r \to \infty$ y, en consecuencia, será una solución posible de (73), a la que corresponderá una función de onda ψ que tiende a cero para $r \to \infty$ según la ley r^{-1}. Así pues, además de los niveles discretos de energía negativa (80) son posibles todos los niveles de energía positiva. Los estados de energía positiva no

son estados ligados, pues para ellos la integral $\int^{\infty}|\chi_0|^2 dr$ no converge en el infinito. (Estos estados corresponden a las órbitas hiperbólicas de la teoría de Bohr.)

40. *Reglas de selección*

Un sistema dinámico abandonado en un estado estacionario, permanecerá en dicho estado mientras no actúen sobre él fuerzas exteriores. Sin embargo, en la práctica, sobre un sistema atómico cualquiera actúan frecuentemente campos electromagnéticos externos, bajo cuya acción puede pasar el sistema de un estado estacionario a otro. La teoría de estas transiciones será desarrollada en los §§ 45 y 46. Una consecuencia de esta teoría es que, con gran aproximación, no pueden ocurrir transiciones entre dos estados por influencia de una radiación electromagnética, si en una representación de Heisenberg en la que dicho par de estados son dos estados básicos, el elemento de matriz del desplazamiento eléctrico total **D** del sistema relativo a dicho par de estados es nulo. Pero ocurre que para muchos sistemas atómicos la mayoría de los elementos de matriz de **D** en una representación de Heisenberg son nulos y, por tanto, existen severas limitaciones de las posibilidades de transición. Las reglas que expresan dichas limitaciones se denominan *reglas de selección.*

La idea de las reglas de selección se puede profundizar mediante la aplicación más detallada de la teoría de §§ 44 y 45, y según ella los elementos de matriz de las distintas componentes cartesianas del vector **D** están asociadas con los distintos estados de polarización de la radiación electromagnética. La naturaleza de dicha asociación es tal que si considerásemos los elementos de matriz o, mejor dicho, sus partes reales como amplitudes de osciladores armónicos que interaccionan con el campo de radiación de acuerdo con las leyes de la electrodinámica clásica, obtendríamos el mismo resultado.

Hay un procedimiento general para obtener todas las reglas de selección, que es el siguiente. Llamemos α a las constantes del movimiento diagonales en la representación de Heisenberg, y sea D una de las componentes cartesianas de **D**. Hemos de obtener una ecuación algebraica que relacione D y las constantes α en la que no intervenga ninguna otra variable dinámica más que la D y las α y que sea lineal respecto a D. Tal ecuación será de la forma

$$\sum_r f_r D g_r = 0, \tag{81}$$

en la que las f_r y las g_r son funciones únicamente de las α. Si expresamos esta ecuación en función de los representantes resulta

$$\sum_r f_r(\alpha')\langle\alpha'|D|\alpha''\rangle g_r(\alpha'') = 0,$$

o bien

$$\langle\alpha'|D|\alpha''\rangle \sum_r f_r(\alpha')g_r(\alpha'') = 0,$$

que nos dice que $\langle\alpha'|D|\alpha''\rangle = 0$ a no ser que

$$\sum_r f_r(\alpha')g_r(\alpha'') = 0. \tag{82}$$

Esta última ecuación que da la relación que tiene que existir entre α' y α'' para que $\langle\alpha'|D|\alpha''\rangle$ no sea nulo, constituye la regla de selección correspondiente a la componente D de $\mathbf{D}$.

El estudio del oscilador armónico hecho en § 34 nos proporciona un ejemplo de regla de selección. La ecuación (8) es de la forma (81) con $\bar{\eta}$ en lugar de D, y donde H juega el papel de las α; de ella se deduce que todos los elementos de matriz $\langle H'|\bar{\eta}|H''\rangle$ de $\bar{\eta}$ son nulos excepto aquellos para los que $H'' - H' = \hbar\omega$. El resultado complejo conjugado de éste es que los elementos de matriz $\langle H'|\eta|H''\rangle$ de η son todos nulos excepto aquellos para los que se verifica $H'' - H' = -\hbar\omega$. Como q es igual a $\bar{\eta} - \bar{\eta}$ multiplicado por un factor numérico, sus elementos de matriz $\langle H'|q|H''\rangle$ serán todos nulos a excepción de los correspondientes a $H'' - H' = \pm\hbar\omega$. Si el oscilador armónico tiene carga eléctrica, su desplazamiento eléctrico D será proporcional a q. Así resulta que la regla de selección permite únicamente transiciones en las que la energía H cambie en un solo cuanto $\hbar\omega$.

Ahora vamos a obtener las reglas de selección de m_z y k para un electrón que se mueva en un campo de fuerzas central. Las componentes del desplazamiento eléctrico ahora son proporcionales a las coordenadas cartesianas x, y, z. Tomemos primero m_z que, como sabemos, conmuta con z

$$m_z z - z m_z = 0.$$

Esta ecuación es del tipo exigido (81), y nos da la regla de selección

$$m_z' - m_z'' = 0$$

para la componente z del desplazamiento. Además, a partir de las ecuaciones (23) resulta

$$[m_z, [m_z, x]] = [m_z, y] = -x$$

o bien,

$$m_z^2 x - 2m_z x m_z + x m_z^2 - \hbar^2 x = 0,$$

que también es del tipo (81), y que nos da la regla de selección

$$m_z'^2 - 2m_z' m_z'' + m_z''^2 - \hbar^2 = 0$$

o bien,

$$(m_z' - m_z'' - \hbar)(m_z' - m_z'' + \hbar) = 0$$

para la componente x del desplazamiento. La regla de selección para la componente y será la misma. Por tanto, las reglas de selección para m_z son

las siguientes: *en las transiciones asociadas a una radiación cuya polarización corresponde a un dipolo eléctrico orientado según el eje z, m'_z no puede cambiar, mientras que en las transiciones asociadas a una polarización correspondiente a un dipolo eléctrico orientado según la dirección del eje x o según la dirección del eje y, m'_z tiene que cambiar en $\pm\hbar$.*

Podemos determinar con más precisión el estado de polarización de la radiación asociada a una transición en la que m'_z cambie en $\pm\hbar$, teniendo en cuenta la condición para que los elementos de matriz de $x+iy$ y de $x-iy$ no sean nulos. Tenemos

$$[m_z, x+iy] = y - ix = -i(x+iy)$$

o sea,

$$m_z(x+iy) - (x+iy)(m_z+\hbar) = 0,$$

que también es una ecuación del tipo (81). De ella deducimos la condición

$$m'_z - m''_z - \hbar = 0$$

que debe verificarse para que $\langle m'_z|x+iy|m''\rangle$ no sea nulo. Análogamente obtenemos

$$m'_z - m''_z + \hbar = 0$$

como condición para que $\langle m'_z|x-iy|m''\rangle$ no sea nulo. Luego

$$\langle m'_z|x-iy|m'_z-\hbar\rangle = 0$$

o sea

$$\langle m'_z|x|m'_z-\hbar\rangle = i\langle m'_z|y|m'_z-\hbar\rangle = (a+ib)e^{i\omega t}$$

en donde a, b y ω son reales. La ecuación compleja conjugada de ésta es

$$\langle m'_z-\hbar|x|m'_z\rangle = -i\langle m'_z-\hbar|y|m'_z\rangle = (a-ib)e^{-i\omega t}$$

Por tanto, el vector $\frac{1}{2}\{\langle m'_z|\mathbf{D}|m'_z-\hbar\rangle + \langle m'_z-\hbar|\mathbf{D}|m'_z\rangle\}$, que determina el estado de polarización de la radiación asociada a las transiciones en las que $m''_z = m'_z-\hbar$, tiene las tres componentes siguientes

$$\left.\begin{aligned}
&\tfrac{1}{2}\{\langle m'_z|x|m'_z-\hbar\rangle + \langle m'_z-\hbar|x|m'_z\rangle\}\\
&\qquad = \tfrac{1}{2}\{(a+ib)e^{i\omega t} + (a-ib)e^{-i\omega t}\} = a\cos\omega t - b\,\mathrm{sen}\,\omega t,\\
&\tfrac{1}{2}\{\langle m'_z|y|m'_z-\hbar\rangle + \langle m'_z-\hbar|y|m'_z\rangle\}\\
&\quad = \tfrac{1}{2}i\{-(a+ib)e^{i\omega t} + (a-ib)e^{-i\omega t}\} = a\,\mathrm{sen}\,\omega t + b\cos\omega t,\\
&\tfrac{1}{2}\{\langle m'_z|z|m'_z-\hbar\rangle + \langle m'_z-\hbar|z|m'_z\rangle\} = 0.
\end{aligned}\right\}\quad(83)$$

Por la forma de estas componentes vemos que la radiación asociada que se propaga en la dirección del eje z estará polarizada circularmente, la que se propaga en una dirección cualquiera del plano xy lo estará rectilí-

neamente en dicho plano, y las que se propagan en direcciones intermedias lo estarán elípticamente. El sentido de la polarización circular de la radiación que se mueve en la dirección del eje z dependerá de si ω es positivo o negativo, que dependerá, a su vez, de cuál de los dos estados m'_z o $m''_z = m'_z - \hbar$ tenga mayor energía.

Vamos a determinar ahora la regla de selección para k. Tenemos

$$\begin{aligned}[k(k+\hbar), z] &= [m_x^2, z] + [m_y^2, z] \\ &= -ym_x - m_x y + xm_y + m_y x \\ &= 2(m_y x - m_x y + i\hbar z) \\ &= 2(m_y x - ym_x) = 2(xm_y - m_x y).\end{aligned}$$

Análogamente

$$[k(k+\hbar), x] = 2(ym_z - m_y z)$$

y

$$[k(k+\hbar), y] = 2(m_x z - xm_z).$$

Luego,

$$\begin{aligned}&[k(k+\hbar), [k(k+\hbar), z]] \\ &\quad = 2[k(k+\hbar), m_y x - m_x y + i\hbar z] \\ &\quad = 2m_y[k(k+\hbar), x] - 2m_x[k(k+\hbar), y] + 2i\hbar[k(k+\hbar), z] \\ &\quad = 4m_y(ym_z - m_y z) - 4m_x(m_x z - xm_z) + 2\{k(k+\hbar)z - zk(k+\hbar)\} \\ &\quad = 4(m_x x + m_y y + m_z z)m_z - 4(m_x^2 + m_y^2 + m_z^2)z + \\ &\qquad\qquad + 2\{k(k+\hbar)z - zk(k+\hbar)\}.\end{aligned}$$

De (22) deducimos

$$m_x x + m_y y + m_z z = 0 \tag{84}$$

luego

$$[k(k+\hbar), [k(k+\hbar), z]] = -2\{k(k+\hbar)z + zk(k+\hbar)\},$$

que da

$$\begin{aligned}k^2(k+\hbar)^2 z - 2k(k+\hbar)zk(k+\hbar) + zk^2(k+\hbar)^2 - \\ - 2\hbar^2\{k(k+\hbar)z + zk(k+\hbar)\} = 0.\end{aligned} \tag{85}$$

Para x e y resultan ecuaciones similares. Dichas ecuaciones son del tipo exigido (81) y de ellas deducimos la regla de selección

$$\begin{aligned}k'^2(k'+\hbar)^2 - 2k'(k'+\hbar)k''(k''+\hbar) + k''^2(k''+\hbar)^2 - \\ - 2\hbar^2 k'(k'+\hbar) - 2\hbar^2 k''(k''+\hbar) = 0,\end{aligned}$$

que se reduce a

$$(k' + k'' + 2\hbar)(k' + k'')(k' - k'' + \hbar)(k' - k'' - \hbar) = 0.$$

Únicamente puede tener lugar una transición entre dos estados k' y k'' si se anula alguno de estos cuatro factores.

Ahora bien, el primero de dichos factores $(k' + k'' + 2\hbar)$ no se puede anular nunca ya que los autovalores de k son todos positivos o nulos. El segundo $(k' + k'')$ sólo se puede anular si $k' = 0$ y $k'' = 0$. Pero no pueden ocurrir transiciones entre estados con dicho valor de k debido a otras reglas

de selección, como podemos ver con el siguiente argumento. Si dos estados que (designamos respectivamente con una y dos primas) son tales que $k' = 0$ y $k'' = 0$, entonces según (41) y los resultados correspondientes para m_x y m_y debe ser $m'_x = m'_y = m'_z = 0$ y $m''_x = m''_y = m''_z = 0$. La regla de selección para m_z nos dice que cuando el valor de m_z no cambia en la transición, los elementos de matriz de x y de y correspondientes a los dos estados son nulos y de la regla de selección correspondiente para m_x y m_y deduciríamos que el elemento de matriz de z también es nulo. Por tanto, no pueden ocurrir transiciones entre tales estados. Así la regla de selección para k se reduce a

$$(k' - k'' + \hbar)(k' - k'' - \hbar) = 0,$$

que muestra que k *tiene que cambiar en* $\pm\hbar$. Dicha regla de selección se puede escribir en la forma

$$k'^2 - 2k'k'' + k''^2 - \hbar^2 = 0,$$

y como ésta es la condición para que el elemento de matriz $\langle k'|z|k''\rangle$ no sea nulo, resulta la ecuación

$$k^2z - 2kzk + zk^2 - \hbar^2 z = 0,$$

o sea

$$[k, [k, z]] = -z, \tag{86}$$

resultado que no es posible obtener fácilmente por un procedimiento más directo.

Como ejemplo final vamos a obtener la regla de selección para la magnitud K del momento angular total $\mathbf{M}$ de un sistema atómico general. Sean x, y, z las coordenadas de uno de los electrones. Hemos de obtener la condición para que los elementos de matriz (K', K'') de x, y, z no sean nulos. Es evidente que dicha condición equivale a la condición de que los elementos de matriz (K', K'') de λ_1, λ_2 y λ_3 no sean nulos, donde λ_1, λ_2 y λ_3 son tres funciones lineales de x, y, z independientes con coeficientes numéricos o, más en general, con coeficientes cualesquiera con tal de que conmuten con K y que, por tanto, estén representados por matrices diagonales respecto a K. Sea

$$\begin{aligned}
\lambda_0 &= M_x x + M_y y + M_z z,\\
\lambda_x &= M_y z - M_z y - i\hbar x,\\
\lambda_y &= M_z x - M_x z - i\hbar y,\\
\lambda_z &= M_x y - M_y x - i\hbar z,
\end{aligned}$$

Según (29), tenemos

$$\begin{aligned}
M_x\lambda_x + M_y\lambda_y + M_z\lambda_z &= \sum_{xyz} (M_xM_yz - M_xM_zy - i\hbar M_x x)\\
&= \sum_{xyz} (M_xM_y - M_yM_x - i\hbar M_z)z = 0
\end{aligned} \tag{87}$$

Por lo tanto, λ_x, λ_y y λ_z no son funciones de x, y, z linealmente independientes. Sin embargo, dos de ellas cualesquiera junto con λ_0 sí son tres funciones de x, y, z linealmente independientes y que, por tanto, se pueden tomar como las anteriores λ_1, λ_2 y λ_3, pues los coeficientes M_x, M_y y M_z conmutan todos ellos con K. Nuestro problema se reduce, por consiguiente, a hallar la condición para que los elementos de matriz (K', K'') de λ_0, λ_x, λ_y y λ_z no sean nulos. El significado físico de las λ es el siguiente: λ_0 es proporcional a la componente del vector (x, y, z) en la dirección del vector $\mathbf{M}$, y λ_x, λ_y y λ_z son proporcionales a las componentes cartesianas de la componente de (x, y, z) perpendicular a $\mathbf{M}$.

Como λ_0 es un escalar tiene que conmutar con K. Resulta, pues, que únicamente pueden ser distintos de cero los elementos de λ_0 diagonales, es decir, $\langle K'|\lambda_0|K'\rangle$, luego, la regla de selección concerniente a λ_0 es que K no puede cambiar en la transición. Aplicando (30) al vector λ_x, λ_y, λ_z, tenemos

$$[M_z, \lambda_x] = \lambda_y, \qquad [M_z, \lambda_y] = -\lambda_x, \qquad [M_z, \lambda_z] = 0.$$

Estas relaciones entre M_z y λ_x, λ_y, λ_z son exactamente de la misma forma que las relaciones (23) y (24) entre m_z y x, y, z, y asimismo (87) es también de la misma forma que (84). Por tanto, las variables dinámicas λ_x, λ_y y λ_z, tienen las mismas propiedades respecto al momento angular $\mathbf{M}$ que tenían las x, y, z respecto a $\mathbf{m}$. En consecuencia, la deducción de la regla de selección para k cuando el desplazamiento eléctrico es proporcional a (x, y, z) puede repetirse para hallar la regla de selección para K cuando el desplazamiento es proporcional a $(\lambda_x, \lambda_y, \lambda_z)$. De este modo encontramos que en lo que concierne a λ_x, λ_y y λ_z, la regla de selección para K es que en la transición ha de cambiar en $\pm\hbar$.

Reuniendo los resultados, obtenemos la regla de selección para K según la cual tiene que cambiar en 0 o en $\pm\hbar$. Hemos considerado únicamente el desplazamiento eléctrico producido por un solo electrón, pero esta misma regla de selección ha de ser válida para cada electrón y en consecuencia ha de valer también para el desplazamiento eléctrico total.

41. El efecto Zeeman en el átomo de hidrógeno

Ahora vamos a considerar el sistema constituido por un átomo de hidrógeno sometido a la acción de un campo magnético uniforme. El hamiltoniano (57) con $V = -e^2/r$ que caracteriza al átomo de hidrógeno en ausencia de campo externo, vendrá modificado por el campo magnético. En mecánica clásica, la modificación consiste en sustituir las componentes del momento p_x, p_y y p_z por $p_x + e/c \cdot A_x$, $p_y + e/c \cdot A_y$, $p_z + e/c \cdot A_z$, en donde A_x, A_y y A_z son las componentes del vector potencial que caracteriza

al campo. En un campo uniforme de magnitud $\mathcal{H}$ según la dirección del eje z, podemos tomar $A_x = -\frac{1}{2}\mathcal{H}y$, $A_y = \frac{1}{2}\mathcal{H}x$, $A_z = 0$. El hamiltoniano clásico sería entonces

$$H = \frac{1}{2m}\left\{\left(p_x - \frac{1}{2}\frac{e}{c}\mathcal{H}y\right)^2 + \left(p_y + \frac{1}{2}\frac{e}{c}\mathcal{H}x\right)^2 + p_z^2\right\} - \frac{e^2}{r}.$$

Este hamiltoniano puede tomarse también en la teoría cuántica añadiéndole el término que da el efecto del spin del electrón. Según la evidencia experimental y según la teoría del capítulo XI, el electrón tiene un momento magnético igual a $-e\hbar/2mc \cdot \boldsymbol{\sigma}$, donde $\boldsymbol{\sigma}$ es el vector de spin de § 37. La energía de este momento magnético en el campo magnético será $e\hbar\mathcal{H}/2mc \cdot \sigma_z$. Luego, el hamiltoniano cuántico total será

$$H = \frac{1}{2m}\left\{\left(p_x - \frac{1}{2}\frac{e}{c}\mathcal{H}y\right)^2 + \left(p_y + \frac{1}{2}\frac{e}{c}\mathcal{H}x\right)^2 + p_z^2\right\} - \frac{e^2}{r} + \frac{e\hbar\mathcal{H}}{2mc}\sigma_z. \quad (88)$$

En rigor deberíamos añadir otros términos en este hamiltoniano que diesen la interacción del momento magnético del electrón con el campo eléctrico del núcleo atómico; pero dicho efecto es pequeño, del mismo orden de magnitud que la corrección que resulta al tener en cuenta la mecánica relativista, y por ahora no lo consideraremos. Ya lo tendremos en cuenta más adelante cuando estudiemos la teoría relativista del electrón en el capítulo XI.

Si el campo magnético no es muy grande, podemos despreciar los términos en $\mathcal{H}^2$, y con ello el hamiltoniano (88) se reduce a

$$\begin{aligned} H &= \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) - \frac{e^2}{r} + \frac{e\mathcal{H}}{2mc}(xp_y - yp_x) + \frac{e\hbar\mathcal{H}}{2mc}\sigma_z \\ &= \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) - \frac{e^2}{r} + \frac{e\mathcal{H}}{2mc}(m_z + \hbar\sigma_z). \end{aligned} \quad (89)$$

Los términos extras debidos al campo magnético son ahora $e\mathcal{H}/2mc \cdot (m_z + \hbar\sigma_z)$. Pero dichos términos conmutan con el hamiltoniano total y, en consecuencia, son constantes del movimiento. Esto hace que el problema sea muy sencillo. Los estados estacionarios del sistema, es decir, los autoestados del hamiltoniano (89) serán los mismos que cuando no había campo magnético alguno, o sea los autoestados comunes a los observables m_z y σ_z o cuando menos al observable $m_z + \hbar\sigma_z$, y los niveles de energía del sistema serán los mismos que en ausencia del campo, dados por (80) si consideramos únicamente estados ligados a los que habrá que añadir un autovalor de $e\mathcal{H}/2mc \cdot (m_z + \hbar\sigma_z)$. Luego, los estados estacionarios del sistema en ausencia de campo para los que m_z tenía el valor numérico m'_z múltiplo

entero de $\hbar$, y σ_z o a su vez el valor numérico $\sigma'_z = \pm 1$, seguirán siendo estacionarios cuando apliquemos el campo. Su energía vendrá aumentada en una cantidad que es la suma de dos partes; una parte $e\mathcal{H}/2mc \cdot m'_z$ que proviene del momento angular orbital y que se puede considerar debida a un momento magnético orbital $-em'_z/2mc$, y otra parte $e\mathcal{H}/2mc \cdot \hbar\sigma'_z$ que es debida al spin. La relación entre el momento magnético orbital y el momento angular orbital m'_z vale $-e/2mc$, que es la mitad de lo que vale la relación entre el momento magnético de spin y el momento angular de spin. A veces se hace referencia a este hecho como anomalía del spin.

Como los niveles de energía dependen ahora de m_z, la regla de selección obtenida en la sección anterior para m_z se puede confrontar directamente con la experiencia. Tomemos una representación de Heisenberg en la que entre otras constantes del movimiento, m_z y σ_z sean diagonales. La regla de selección para m_z exige en este caso que m_z cambie en $\hbar$, 0, o $-\hbar$, mientras que σ_z por conmutar con el desplazamiento eléctrico no deberá cambiar. Luego, la variación de energía en la transición entre los dos estados diferirá del valor en ausencia de campo electromagnético en una cantidad igual a $e\hbar\mathcal{H}/2mc$, 0, o $-e\hbar\mathcal{H}/2mc$. Por tanto, según la condición de Bohr para las frecuencias, la frecuencia de la radiación electromagnética asociada diferirá de la correspondiente a la ausencia de campo magnético en $e\mathcal{H}/4\pi mc$, 0, o $-e\mathcal{H}/4\pi mc$. Esto significa que cada una de las líneas espectrales que aparecía en ausencia de campo magnético se desdoblará en tres componentes. Si consideramos la radiación que se emite en la dirección del eje z, según (83), las dos componentes extremas estarán polarizadas circularmente, mientras que la línea central no desplazada será de intensidad nula. Estos resultados están de acuerdo con la experiencia y también con la teoría clásica del efecto Zeeman.

VII

TEORÍA DE PERTURBACIONES

42. *Generalidades*

En la sección anterior hemos estudiado rigurosamente algunos sistemas sencillos de la teoría cuántica. Sin embargo, la mayoría de los problemas que se plantean en esta teoría no se pueden resolver exactamente con los recursos matemáticos actuales, pues conducen a ecuaciones cuyas soluciones no se pueden expresar mediante un número finito de términos con las funciones ordinarias del análisis. Para dichos problemas puede utilizarse, con frecuencia, un método de perturbaciones. Dicho método consiste en separar el hamiltoniano en dos partes, de las cuales una ha de ser sencilla y la otra pequeña. La primera de ellas puede ser considerada como el hamiltoniano de un sistema simplificado o no perturbado que se puede estudiar con todo rigor, mientras que la adición de la segunda no introducirá más que pequeñas correcciones de la naturaleza de una perturbación en la solución del sistema no perturbado. La exigencia de que la primera parte sea sencilla implica en la práctica que no sea función explícita del tiempo. Si la segunda parte contiene un pequeño factor numérico ε, podemos obtener la solución de las ecuaciones del sistema perturbado en la forma de una serie de potencias de ε, que si converge nos dará la solución del problema con tanta aproximación como queramos. Incluso cuando dicha serie no converge, ocurre muchas veces que la primera aproximación obtenida con ella es bastante precisa.

En la teoría de perturbaciones hay dos métodos. En uno se considera que la perturbación provoca *una modificación de los estados de movimiento* del sistema no perturbado. En el otro en vez de considerar una modificación de los estados del sistema no perturbado, se supone que el sistema perturbado en lugar de permanecer siempre en uno de dichos estados, cambia constantemente de uno a otro, o sea, se considera que bajo la influencia de la perturbación se *producen transiciones*. El método que debe emplearse en cada caso particular depende de la naturaleza del problema que ha de resolverse. El primer método, por lo general, sólo tiene utilidad cuando la energía de la perturbación (es decir, la corrección que debe añadirse al

hamiltoniano del sistema no perturbado), no es función explícita del tiempo, en cuyo caso se aplica a los estados estacionarios. También se puede utilizar para calcular resultados que no se refieren a un instante de tiempo particular, tales como niveles de energía o estados estacionarios del sistema perturbado, y en los problemas de colisión, para calcular la probabilidad de que las partículas sean dispersadas bajo un cierto ángulo. En cambio, en todos los problemas en que entra en consideración el tiempo, como los que hacen referencia a los fenómenos transitorios que se producen cuando se aplica repentinamente una perturbación, o en general todos aquellos problemas en los que la perturbación varía de alguna manera con el tiempo (es decir, cuando la energía de la perturbación es función explícita del tiempo), tenemos que emplear el segundo método. Asimismo hemos de emplear el segundo método en los problemas de colisión para calcular las probabilidades de absorción y de emisión aunque la energía de perturbación no sea función explícita del tiempo, pues contrariamente a lo que ocurre con las probabilidades de dispersión, estas probabilidades no se pueden definir sin hacer referencia a un estado de cosas que varía con el tiempo.

Podemos resumir los rasgos característicos de ambos métodos de la siguiente forma: en el primer método comparamos los estados estacionarios del sistema perturbado con los del sistema no perturbado, y en el segundo método se considera un estado estacionario del sistema no perturbado y se estudia cómo varía bajo la influencia de la perturbación.

43. *Variación de los niveles de energía producida por una perturbación*

Vamos a aplicar el primero de los dos métodos mencionados para calcular las variaciones de los niveles de energía de un sistema producidas por una perturbación. Supondremos que la energía de perturbación, al igual que el hamiltoniano del sistema no perturbado, no es función explícita del tiempo. Desde luego, nuestro problema sólo tiene sentido si los niveles de energía son discretos, y si las diferencias entre ellos son grandes en comparación con los cambios que la perturbación introduce en ellos. Este hecho hace que el planteamiento de los problemas de perturbación por el primer método sea distinto cuando los niveles de energía del sistema no perturbado son discretos y cuando son continuos.

Sea

$$H = E + V, \tag{1}$$

el hamiltoniano del sistema perturbado, en donde E representa el hamiltoniano del sistema no perturbado y V la pequeña energía de perturbación. Por hipótesis, cada autovalor H' de H está muy próximo a uno y sólo un autovalor E' de E. Utilizaremos el mismo número de primas para de-

signar un autovalor de H y el autovalor de E que está muy próximo a él. Por tanto, H'' diferirá de E'' en una pequeña cantidad de orden V y, en cambio, diferirá de E' en una cantidad que no es pequeña salvo si $E' = E''$. Deberemos utilizar siempre cuidadosamente distinto número de primas para designar autovalores de H y de E que no nos interese hacer constar que están muy próximos.

Para obtener los autovalores de H hemos de resolver la ecuación

$$H|H'\rangle = H'|H'\rangle$$

o bien

$$(H' - E)|H'\rangle = V|H'\rangle. \tag{2}$$

Sea $|0\rangle$ un autoket de E perteneciente al autovalor E' y supongamos que los $|H'\rangle$ y H' que verifican (2) difieren de $|0\rangle$ y E' en pequeñas cantidades que podemos expresar así

$$\left.\begin{aligned} |H'\rangle &= |0\rangle + |1\rangle + |2\rangle + \ldots, \\ H' &= E' + a_1 + a_2 + \ldots, \end{aligned}\right\} \tag{3}$$

siendo $|1\rangle$ y a_1 infinitésimos de primer orden (es decir, del mismo orden que V), $|2\rangle$ y a_2 de segundo orden, etc. Sustituyendo estas expresiones en (2) resulta

$$\{E' - E + a_1 + a_2 + \ldots\}\{|0\rangle + |1\rangle + |2\rangle + \ldots\} = V\{|0\rangle + |1\rangle \ldots\}.$$

Si ahora separamos los términos de orden cero, los de primer orden, los de segundo y así sucesivamente, resultan las siguientes ecuaciones

$$\left.\begin{aligned} (E' - E)|0\rangle &= 0, \\ (E' - E)|1\rangle + a_1|0\rangle &= V|0\rangle, \\ (E' - E)|2\rangle + a_1|1\rangle + a_2|0\rangle &= V|1\rangle, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{aligned}\right\} \tag{4}$$

La primera de estas ecuaciones no nos dice más que lo que habíamos supuesto previamente, o sea que $|0\rangle$ es un autoket de E perteneciente al autovalor E'. Las otras nos permiten calcular las distintas correcciones $|1\rangle$, $|2\rangle$, ..., y a_1, a_2, ...

Para seguir discutiendo estas ecuaciones es conveniente introducir una representación en la que E sea diagonal, es decir, una representación de Heisenberg para el sistema no perturbado con E como uno de los observables con cuyos autovalores se designan los representantes. Representemos por β los demás observables en el caso de que sean necesarios, como ocurre cuando existe más de un autoestado de E perteneciente a un mismo autovalor. Así los bras básicos serán $\langle E''\beta''|$. Como $|0\rangle$ es un autoket de E perteneciente al autovalor E', tenemos

$$\langle E''\beta''|0\rangle = \delta_{E''E'} f(\beta''), \tag{5}$$

siendo $f(\beta'')$ una cierta función de las variables β''. Con ayuda de este resultado, la segunda de las ecuaciones (4) escrita en forma de representantes será

$$(E' - E'')\langle E''\beta''|1\rangle + a_1 \delta_{E''E'} f(\beta'') = \sum_{\beta'} \langle E''\beta''|V|E'\beta'\rangle f(\beta'). \tag{6}$$

Haciendo $E'' = E'$, queda

$$a_1 f(\beta'') = \sum_{\beta'} \langle E'\beta''|V|E'\beta'\rangle f(\beta'). \tag{7}$$

En lo que concierne a las variables β', la ecuación (7) es del tipo clásico en la teoría de autovalores. En ella vemos que los diversos valores posibles de a_1 son los autovalores de la matriz $\langle E'\beta''|V|E'\beta'\rangle$. Esta matriz es una parte del representante de la energía de perturbación en la representación de Heisenberg del sistema no perturbado, a saber, la parte de matriz que corresponde a los elementos que hacen referencia al nivel de energía E' tanto respecto a filas como respecto a columnas. Cada uno de estos valores de a_1 da un nivel de energía del sistema perturbado próximo al nivel de energía E' del sistema no perturbado † en la aproximación de primer orden. Por tanto, pueden existir distintos niveles de energía del sistema no perturbado independientes que pertenezcan al nivel de energía E'. La perturbación puede producir un desdoblamiento total o parcial de los niveles de energía que para el sistema no perturbado tienen todos el mismo valor E'.

La ecuación (7) determina asimismo los representantes $\langle E''\beta''|0\rangle$ de los estados estacionarios del sistema perturbado pertenecientes a los niveles de energía próximos a E', en la aproximación de orden cero, ya que cada solución $f(\beta')$ de (7) sustituida en (5) constituye uno de dichos representantes. Cada uno de dichos estados estacionarios del sistema perturbado está próximo a un estado estacionario del sistema no perturbado, pero el recíproco que consistiría en que cada estado estacionario del sistema no perturbado estuviera próximo a un estado estacionario del sistema perturbado no es cierto, ya que el estado estacionario general del sistema no perturbado perteneciente al nivel de energía E' está representado por el segundo miembro de (5) en donde $f(\beta'')$ es una función arbitraria. El problema de hallar los estados estacionarios del sistema no perturbado que aproximan a estados estacionarios del sistema perturbado, o sea, el problema de hallar las funciones $f(\beta')$ de la ecuación (7) corresponde al cálculo de las 'perturbaciones seculares' de la mecánica clásica. Observemos que

---

† Para distinguir los niveles de energía que así resultan necesitaríamos una notación más elaborada, ya que de acuerdo con nuestra actual notación todos ellos tendrían que tener un número de primas igual al que sirve para designar el nivel de energía del sistema no perturbado del que derivan. Sin embargo, para nuestros propósitos no será necesaria tal notación.

los resultados anteriores no dependen de los valores de los elementos de matriz de la energía de perturbación que hacen referencia a dos niveles de energía distintos del sistema no perturbado.

Veamos cómo se simplifican los resultados establecidos aquí en el caso particularmente sencillo de que no exista más que un único estado estacionario del sistema no perturbado perteneciente a cada nivel de energía.† En este caso E determina la representación por sí sólo, y no se precisa ningún β. La suma de (7) se reduce ahora a un único término, y así resulta

$$a_1 = \langle E'|V|E'\rangle. \tag{8}$$

No existe más que un único nivel de energía del sistema perturbado próximo a cada nivel de energía del sistema no perturbado, y *la variación de energía en la aproximación de primer orden es igual al correspondiente elemento diagonal de la energía de perturbación en la representación de Heisenberg del sistema no perturbado, o sea, al valor medio de la energía de perturbación en el correspondiente estado no perturbado.* La última formulación del resultado es la misma que en mecánica clásica cuando el sistema no perturbado es múltiplemente periódico.

Vamos a calcular la corrección de segundo orden a_1 de los niveles de energía en el caso de que el sistema no perturbado no sea degenerado. La ecuación (5) se convierte en este caso en

$$\langle E''|0\rangle = \delta_{E''E'}$$

prescindiendo de un factor numérico sin importancia, y la ecuación (6) es ahora

$$(E' - E'')\langle E''|1\rangle + a_1\,\delta_{E''E'} = \langle E''|V|E'\rangle.$$

De ella podemos despejar el valor de $\langle E''|1\rangle$ cuando $E'' \neq E'$,

$$\langle E''|1\rangle = \frac{\langle E''|V|E'\rangle}{E' - E''}. \tag{9}$$

La tercera de las ecuaciones (4) escrita en forma de representantes es

$$(E' - E'')\langle E''|2\rangle + a_1\langle E''|1\rangle + a_2\,\delta_{E''E'} = \sum_{E'''} \langle E''|V|E'''\rangle\langle E'''|1\rangle.$$

Y haciendo $E'' = E'$ queda

$$a_1\langle E'|1\rangle + a_2 = \sum_{E'''} \langle E'|V|E'''\rangle\langle E'''|1\rangle,$$

---

† Los sistemas con un único estado estacionario para cada nivel de energía se denominan a menudo sistemas *no degenerados,* y los que poseen dos o más estados estacionarios pertenecientes al mismo nivel de energía, sistemas *degenerados;* sin embargo, esta nomenclatura no es muy apropiada desde el punto de vista moderno.

que con ayuda de (8) se reduce a

$$a_2 = \sum_{E''\neq E'} \langle E'|V|E''\rangle\langle E''|1\rangle.$$

Sustituyendo el valor $\langle E''|1\rangle$ dado por (9) resulta finalmente

$$a_2 = \sum_{E''\neq E'} \frac{\langle E'|V|E''\rangle\langle E''|V|E'\rangle}{E' - E''}$$

que da como corrección total para la energía en segunda aproximación el valor

$$a_1 + a_2 = \langle E'|V|E'\rangle + \sum_{E''\neq E'} \frac{\langle E'|V|E''\rangle\langle E''|V|E'\rangle}{E' - E''} \tag{10}$$

Este método se puede desarrollar para calcular aproximaciones de órdenes superiores si fuera necesario. Born, Heisenberg y Jordan† han establecido fórmulas de recurrencia generales que dan la corrección correspondiente a la aproximación de orden n en función de las correcciones de orden inferior.

44. *La perturbación considerada como causa de transiciones*

Vamos a considerar ahora el segundo método de calcular perturbaciones que mencionábamos en § 42. Supongamos como antes que tenemos un sistema no perturbado caracterizado por un hamiltoniano E que no es función explícita del tiempo, y una energía de perturbación V que ahora puede ser una función cualquiera del tiempo. El hamiltoniano del sistema perturbado será como antes $H = E + V$. Para este método no significa ninguna diferencia esencial el hecho de que los niveles de energía del sistema no perturbado, es decir, los autovalores de E, constituyan un conjunto discreto o continuo. Pero para concretar tomaremos el caso en que los niveles de energía son discretos. También ahora trabajaremos en una representación de Heisenberg del sistema no perturbado, pero como ahora no representa ninguna ventaja elegir E como uno de los observables cuyos autovalores sirven para designar los representantes, tomaremos un conjunto de α general.

Supongamos que en el instante inicial t_0 el sistema está en un estado en el que las α tienen ciertamente los valores α'. El ket correspondiente a dicho estado es el ket básico $|\alpha'\rangle$. Si no hubiese ninguna perturbación, es decir, si el hamiltoniano fuese E, dicho estado sería estacionario. La perturbación hace que el estado cambie. En el instante t el ket que corresponde

† *Z. f. Physik*, 35 (1925), 565.

al estado en la imagen de Schrödinger será $T|\alpha'\rangle$, según la ecuación (1) de § 27. La probabilidad de que los α tengan los valores α'' será

$$P(\alpha'\alpha'') = |\langle\alpha''|T|\alpha'\rangle|^2. \tag{11}$$

Para $\alpha'' \neq \alpha'$, $P(\alpha'\alpha'')$ es la probabilidad de que haya una transición desde el estado α' al estado α'' durante el intervalo de tiempo de t_0 a t, mientras que $P(\alpha'\alpha')$ es la probabilidad de que no haya ninguna transición. La suma de $P(\alpha'\alpha'')$ para todo α'', como siempre, es igual a uno.

Ahora supongamos que el sistema en lugar de estar inicialmente en el estado α' con certeza, esté en uno y otro de los distintos estados α' con la probabilidad $P_{\alpha'}$. La densidad de Gibbs que corresponde a dicha distribución, según (68) de § 33 es

$$\rho = \sum_{\alpha'} |\alpha'\rangle P_{\alpha'} \langle\alpha'|. \tag{12}$$

En el instante t, cada ket $|\alpha'\rangle$ habrá pasado a $T|\alpha'\rangle$ y cada bra $\langle\alpha'|$ a $\langle\alpha'|\bar{T}$, con lo que ρ se habrá transformado en

$$\rho_t = \sum_{\alpha'} T|\alpha'\rangle P_{\alpha'} \langle\alpha'|\bar{T}. \tag{13}$$

Según (73) de § 33, la probabilidad de que los α tengan los valores α'' será entonces

$$\begin{aligned} \langle\alpha''|\rho_t|\alpha''\rangle &= \sum_{\alpha'} \langle\alpha''|T|\alpha'\rangle P_{\alpha'} \langle\alpha'|\bar{T}|\alpha''\rangle \\ &= \sum_{\alpha'} P_{\alpha'} P(\alpha'\alpha'') \end{aligned} \tag{14}$$

tras aplicar (11). Este resultado indica que la probabilidad de que el sistema esté en el estado α'' en el instante t es igual a la suma de las probabilidades de que en el instante inicial estuviera en un estado cualquiera $\alpha' \neq \alpha''$ y haya hecho una transición desde dicho estado α' al estado α'', más la probabilidad de que el sistema estuviera inicialmente en el estado α'' y no haya hecho ninguna transición. Por tanto, las diversas probabilidades de transición son independientes unas de otras, de acuerdo con las leyes ordinarias de las probabilidades.

Así pues, el problema de calcular las transiciones se reduce a la determinación de las amplitudes de probabilidad $\langle\alpha''|T|\alpha'\rangle$. Éstas se pueden estudiar a partir de la ecuación diferencial en T, (6) de § 27

$$i\hbar\, dT/dt = HT = (E + V)T. \tag{15}$$

Para simplificar el cálculo introducimos el operador

$$T^* = e^{iE(t-t_0)/\hbar} T. \tag{16}$$

Tenemos

$$\begin{aligned} i\hbar\, dT^*/dt &= e^{iE(t-t_0)/\hbar}(-ET + i\hbar\, dT/dt) \\ &= e^{iE(t-t_0)/\hbar} VT = V^*T^*, \end{aligned} \tag{17}$$

en donde

$$V^* = e^{iE(t-t_0)/\hbar} V e^{-iE(t-t_0)/\hbar}, \tag{18}$$

es decir, V^* es el resultado de aplicar una cierta transformación unitaria a V. La ecuación (17) es más apta para el cálculo que (15), pues en (17) la variación de T^* depende enteramente de la perturbación V, y para $V = 0$, T^* conserva el valor inicial de T^* que es igual a la unidad. Según (16) tenemos

$$\langle \alpha'' | T^* | \alpha' \rangle = e^{iE''(t-t_0)/\hbar} \langle \alpha'' | T | \alpha' \rangle,$$

o sea,

$$P(\alpha'\alpha'') = | \langle \alpha'' | T^* | \alpha' \rangle |^2, \tag{19}$$

por lo que T^* y T son igualmente útiles para determinar las probabilidades de transición.

Hasta aquí todo el estudio realizado es exacto. Ahora vamos a suponer que V es un infinitésimo de primer orden y expresaremos T^* en la forma

$$T^* = 1 + T_1^* + T_2^* + \ldots, \tag{20}$$

donde T_1^* será de primer orden, T_2^* de segundo, etc. Sustituyendo (20) en (17) e igualando los términos del mismo orden resulta

$$\left.\begin{aligned} i\hbar\, dT_1^*/dt &= V^*, \\ i\hbar\, dT_2^*/dt &= V^* T_1^*, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{aligned}\right\} \tag{21}$$

De la primera de estas ecuaciones obtenemos

$$T_1^* = -i\hbar^{-1} \int_{t_0}^{t} V^*(t')\, dt', \tag{22}$$

y de la segunda resulta

$$T_2^* = -\hbar^{-2} \int_{t_0}^{t} V^*(t')\, dt' \int_{t_0}^{t'} V^*(t'')\, dt'', \tag{23}$$

y así sucesivamente. Para muchos problemas prácticos es suficiente conservar el término T_1^*, que nos da un valor de la probabilidad de transición $P(\alpha'\alpha'')$ con $\alpha'' \neq \alpha'$

$$\left.\begin{aligned} P(\alpha'\alpha'') &= \hbar^{-2} \left| \langle \alpha'' | \int_{t_0}^{t} V^*(t')\, dt' | \alpha' \rangle \right|^2 \\ &= \hbar^{-2} \left| \int_{t_0}^{t} \langle \alpha'' | V^*(t') | \alpha' \rangle\, dt' \right|^2 \end{aligned}\right\} \tag{24}$$

Así obtenemos la probabilidad de transición en segundo orden de aproxi-

mación. El resultado sólo depende del elemento de matriz $\langle\alpha''|V^*(t')|\alpha'\rangle$ de $V^*(t')$ relativo a los dos estados en cuestión, en el tiempo t' que media entre t_0 y t. Como V^* es real igual que V, tenemos

$$\langle\alpha''|V^*(t')|\alpha'\rangle = \overline{\langle\alpha'|V^*(t')|\alpha''\rangle}$$

luego,

$$P(\alpha'\alpha'') = P(\alpha''\alpha') \tag{25}$$

para el segundo orden de aproximación.

Algunas veces nos interesan transiciones $\alpha' \to \alpha''$ para las que el elemento $\langle\alpha''|V^*|\alpha'\rangle$ es nulo o bien es pequeño comparado con los otros elementos de la matriz de V^*. En estos casos es necesario trabajar con un orden de aproximación mayor. Si solamente conservamos los términos T_1^* y T_2^*, para $\alpha'' \neq \alpha'$ resulta

$$P(\alpha'\alpha'') = \hbar^{-2}\left|\int_{t_0}^{t} \langle\alpha''|V^*(t')|\alpha'\rangle\, dt' - \right.$$

$$\left. - i\hbar^{-1} \sum_{\alpha''' \neq \alpha',\, \alpha''} \int_{t_0}^{t} \langle\alpha''|V^*(t')|\alpha'''\rangle\, dt' \int_{t_0}^{t'} \langle\alpha'''|V^*(t'')|\alpha'\rangle\, dt''\right|^2. \tag{26}$$

Los términos correspondientes a $\alpha''' = \alpha'$ y $\alpha''' = \alpha''$ los hemos omitido en la suma porque teniendo en cuenta que $\langle\alpha''|V^*|\alpha'\rangle$ es pequeño, resultan pequeños comparados con los demás de la suma. A fin de interpretar el resultado (26) podemos suponer que el término

$$\int_{t_0}^{t} \langle\alpha''|V^*(t')|\alpha'\rangle\, dt' \tag{27}$$

da cuenta de una transición directa desde el estado α' al estado α'', mientras que el término

$$- i\hbar^{-1} \int_{t_0}^{t} \langle\alpha''|V^*(t')|\alpha'''\rangle\, dt' \int_{t_0}^{t'} \langle\alpha'''|V^*(t'')|\alpha'\rangle\, dt'' \tag{28}$$

da cuenta de una transición desde el estado α' al estado α''' seguida de una transición desde dicho estado α''' al estado α''. En esta interpretación el estado α''' se denomina *estado intermedio*. Tenemos que sumar el término (27) con los diversos términos (28) correspondientes a los distintos estados intermedios, y tomar el cuadrado del módulo de dicha suma, lo que nos dice que hay interferencia entre los distintos procedimientos de transición — la transición directa y todas las que se verifican a través de estados intermedios — y por tanto, no podemos asignar ningún significado a la

probabilidad de que ocurra cada uno de dichos sucesos por separado. Sin embargo, existe una amplitud de probabilidad para cada uno de ellos. Estudiando el método de perturbaciones con órdenes de aproximación más elevados, llegamos a resultados que se pueden interpretar análogamente mediante procesos de transición más complicados que llevan consigo una sucesión de estados intermedios.

45. *Aplicación a la radiación*

En la sección anterior hemos desarrollado una teoría general de perturbaciones de un sistema atómico, en la que la energía de perturbación podía variar con el tiempo de forma totalmente arbitraria. En la práctica podemos producir una perturbación de este tipo haciendo incidir sobre el sistema energía electromagnética. Veamos a qué se reduce el resultado (24) en este caso.

Si despreciamos los efectos del campo magnético de la radiación incidente, y si además, suponemos que las longitudes de onda de las componentes armónicas de dicha radiación son grandes en comparación con las dimensiones del sistema atómico, entonces la energía de perturbación es simplemente el producto escalar

$$V = (\mathbf{D}, \boldsymbol{\mathcal{E}}), \tag{29}$$

siendo $\mathbf{D}$ el desplazamiento eléctrico total del sistema y $\boldsymbol{\mathcal{E}}$ el campo eléctrico de la radiación incidente. Supongamos que $\boldsymbol{\mathcal{E}}$ es una función del tiempo conocida. Si para simplificar suponemos que la radiación incidente está polarizada rectilíneamente con su vector campo eléctrico orientado en una determinada dirección, llamando D a la componente cartesiana de $\mathbf{D}$ en dicha dirección, la expresión (29) de V se reduce entonces al producto ordinario

$$V = D\mathcal{E},$$

siendo $\mathcal{E}$ la magnitud del vector $\boldsymbol{\mathcal{E}}$. Los elementos de matriz de V serán entonces

$$\langle \alpha'' | V | \alpha' \rangle = \langle \alpha'' | D | \alpha' \rangle \mathcal{E},$$

ya que $\mathcal{E}$ es un número. El elemento de matriz $\langle \alpha'' | D | \alpha' \rangle$ es independiente del tiempo t. Según (18),

$$\langle \alpha'' | V^*(t) | \alpha' \rangle = \langle \alpha'' | D | \alpha' \rangle e^{i(E''-E')(t-t_0)/\hbar} \mathcal{E}(t),$$

con lo que la expresión (24) para la probabilidad de transición nos da

$$P(\alpha'\alpha'') = \hbar^{-2} | \langle \alpha'' | D | \alpha' \rangle |^2 \left| \int_{t_0}^{t} e^{i(E''-E')(t'-t_0)/\hbar} \mathcal{E}(t')\, dt' \right|^2 \tag{30}$$

Si analizamos la radiación incidente durante el intervalo de tiempo de t_0 a t en sus componentes de Fourier, la energía que atraviesa la unidad de área por unidad de intervalo de frecuencia para la frecuencia ν, según la electrodinámica clásica será

$$E_\nu = \frac{c}{2\pi} \left| \int_{t_0}^{t} e^{2\pi i\nu(t'-t_0)} \mathcal{E}(t')\, dt' \right|^2 \tag{31}$$

Comparando este valor con (30), obtenemos

$$P(\alpha'\alpha'') = 2\pi c^{-1}\hbar^{-2} |\langle \alpha'' | D | \alpha' \rangle|^2 E_\nu, \tag{32}$$

siendo

$$\nu = |E'' - E'|/h. \tag{33}$$

Según este resultado vemos que en primer lugar la probabilidad de transición únicamente depende de la componente de Fourier de la radiación incidente cuya frecuencia ν está relacionada con el cambio de energía por (33). (33) nos da la *condición de Bohr para la frecuencia* y nos indica en qué forma pueden ser incluidas en la mecánica cuántica las ideas de Bohr, precursor de la misma.

La teoría elemental que hemos expuesto no nos indica nada acerca de la energía del campo de radiación. Sin embargo, sería lógico suponer que la energía absorbida o liberada por el sistema atómico en el proceso de transición procede de la componente de la radiación de frecuencia ν dada por (33) o va a parar a ella. Esta hipótesis quedará justificada con una teoría de la radiación más completa que expondremos en el capítulo X. A la luz de ella, el resultado (32) debe ser interpretado como la probabilidad de que, si el sistema estaba inicialmente en el estado de menor energía, absorba radiación y pase al estado de mayor energía, y si inicialmente se encontraba en dicho estado, sea *estimulado* por la energía de radiación a emitir y pasar al estado de menor energía. La teoría aquí expuesta tampoco explica el hecho experimental de que cuando el sistema está en el estado de energía superior puede emitir radiación espontáneamente en ausencia de radiación incidente y pasar al estado de energía inferior, pero esto también queda explicado por la teoría de la radiación más completa del capítulo X.

La existencia del fenómeno de emisión estimulada fue inferida por Einstein † mucho antes del descubrimiento de la mecánica cuántica, a partir de una consideración de equilibrio estadístico entre los átomos y el campo de radiación del cuerpo negro que verificaba la ley de Planck. Einstein demostró que la probabilidad de transición para la emisión estimulada tenía

† EINSTEIN, *Phys. Zeits.*, 18 (1917), 121.

que ser igual a la de absorción entre el mismo par de estados, resultado que está de acuerdo con la teoría cuántica aquí expuesta, y asimismo dedujo una relación entre dicha probabilidad de transición y la de emisión espontánea, que está de acuerdo con la teoría del capítulo X.

El elemento de matriz $\langle\alpha''|D|\alpha'\rangle$ de (32) hace el papel de la amplitud de una de las componentes de Fourier de D en la teoría clásica de un sistema múltiplemente periódico en interacción con la radiación. De hecho, la idea de sustituir las componentes de Fourier clásicas por elementos de matriz fue lo que condujo a Heisenberg al descubrimiento de la mecánica cuántica en 1925. Heisenberg supuso que las fórmulas que describen la interacción de un sistema con la radiación en mecánica cuántica podían ser obtenidas sustituyendo las componentes de Fourier del desplazamiento eléctrico total del sistema que figura en las fórmulas clásicas, por los correspondientes elementos de matriz. Si aplicamos esta hipótesis a la emisión espontánea, un sistema con un momento eléctrico total $\mathbf{D}$ que estuviera en el estado α', tendría que emitir radiación de frecuencia $\nu = (E' - E'')/\hbar$, siendo E'' un nivel de energía correspondiente a un estado α'' tal que $E'' < E'$, en la proporción

$$\frac{4}{3}\frac{(2\pi\nu)^4}{c^3}\,|\langle\alpha''|\mathbf{D}|\alpha'\rangle|^2. \tag{34}$$

por unidad de tiempo. La distribución de dicha radiación según las distintas direcciones de emisión y el estado de polarización para cada dirección serían iguales que para un dipolo eléctrico clásico de momento igual a la parte real de $\langle\alpha''|\mathbf{D}|\alpha'\rangle$. Para interpretar esta densidad de emisión de energía radiante por unidad de tiempo como una probabilidad de transición tenemos que dividirla por el cuanto de energía correspondiente a dicha frecuencia, o sea por $h\nu$, y el resultado será la probabilidad de que dicho cuanto sea emitido espontáneamente en la unidad de tiempo, pasando simultáneamente el sistema al estado α'' de menos energía. Estas hipótesis quedan justificadas con la teoría de la radiación que hemos expuesto aquí completada con la teoría de las transiciones espontáneas del capítulo X.

46. *Transiciones producidas por una perturbación independiente del tiempo*

El método de perturbaciones de § 44 sigue siendo válido cuando la energía V no es función explícita del tiempo t. Dado que en este caso el hamiltoniano total H no sería función explícita de t, podríamos emplear si quisiéramos el método de § 43 y hallar sus estados estacionarios. La utilidad de este método dependerá de lo que queramos conocer del sistema. Si hemos de calcular algo que hace referencia explícita al tiempo, como, por ejemplo, la probabilidad de que el sistema esté en un cierto estado

en un instante sabiendo que en otro instante está en otro estado, será más adecuado al método de § 44.

Veamos en qué se convierte el resultado (24) para la probabilidad de transición cuando V no es función explícita de t, tomando $t_0 = 0$ para simplificar la notación. Ahora el elemento de matriz $\langle\alpha''|V|\alpha'\rangle$ es independiente de t, y según (18)

$$\langle\alpha''|V^*(t')|\alpha'\rangle = \langle\alpha''|V|\alpha'\rangle e^{i(E''-E')t'/\hbar}, \tag{35}$$

de modo que

$$\int_0^t \langle\alpha''|V^*(t')|\alpha'\rangle\, dt' = \langle\alpha''|V|\alpha'\rangle \frac{e^{i(E''-E')t/\hbar} - 1}{i(E''-E')/\hbar},$$

siempre que $E'' \neq E'$. Luego, la probabilidad de transición (24) será

$$\begin{aligned} P(\alpha'\alpha'') &= |\langle\alpha''|V|\alpha'\rangle|^2[e^{i(E''-E')t/\hbar} - 1][e^{-i(E''-E')t/\hbar} - 1]/(E''-E')^2 \\ &= 2|\langle\alpha''|V|\alpha'\rangle|^2[1 - \cos\{(E''-E')t/\hbar\}]/(E''-E')^2. \end{aligned} \tag{36}$$

Si E'' es muy diferente de E' dicha probabilidad de transición es pequeña y se mantiene pequeña para todos los valores de t. Este resultado viene impuesto por la ley de la conservación de la energía. La energía total H es constante, y por tanto, la energía propia E (es decir, la energía que se obtiene despreciando la parte V debida a la perturbación) por ser aproximadamente igual a H, ha de ser aproximadamente constante. Esto quiere decir que si E tenía inicialmente el valor numérico E', únicamente existirá una pequeña probabilidad de que en un instante posterior tenga un valor numérico considerablemente distinto de E'.

En cambio, cuando el estado inicial α' es tal que existe otro estado α'' con la misma o casi la misma energía propia E, la probabilidad de una transición al estado final α'' puede ser bastante grande. El caso que tiene interés físico es aquel en que existe un conjunto continuo de estados finales α'' que poseen niveles de energía propia E'' que contienen el nivel de energía propia E' del estado inicial. Dicho estado inicial no tiene por qué formar parte de la serie de estados del conjunto continuo, sino que puede ser un estado discreto aislado o también un estado perteneciente a otro conjunto de estados continuos. Recordando las reglas de § 18 para la interpretación de amplitudes de probabilidad referentes a conjuntos continuos de estados, si $P(\alpha'\alpha'')$ tiene el valor (36), la probabilidad de que haya una transición a un estado final comprendido en el pequeño intervalo de α'' a $\alpha' + d\alpha''$ será $P(\alpha'\alpha'')d\alpha''$ cuando el estado incial α' sea discreto, y será proporcional a dicho valor cuando α' pertenezca a un conjunto continuo.

Supongamos que los α que nos sirven para designar el estado final son E y un conjunto de otras variables dinámicas β, y así tendremos una representación análoga a la de § 43 en el caso en que hay degeneración. (Sin embargo, las β no tienen por qué tener significado para el estado inicial α'.)

Para concretar supondremos que las β sólo tienen autovalores discretos. La probabilidad total de que haya una transición a un estado final α'' para el cual las β tienen los valores β'' y E tiene un valor cualquiera (habrá una gran probabilidad cuando el estado final tenga una energía propia muy próxima a la E' del estado inicial) será (o bien será proporcional a)

$$\int P(\alpha'\alpha'')\, dE''$$
$$= 2\int_{-\infty}^{\infty} |\langle E''\beta''|V|\alpha'\rangle|^2[1-\cos\{(E''-E')t/\hbar\}]/(E''-E')^2\, dE'' \quad (37)$$
$$= 2t\hbar^{-1}\int_{-\infty}^{\infty} |\langle E'+\hbar x/t, \beta''|V|\alpha'\rangle|^2[1-\cos x]/x^2\, dx$$

si hacemos la sustitución $(E''-E')t/\hbar = x$. Para grandes valores de t se reduce a

$$2t\hbar^{-1}|\langle E'\beta''|V|\alpha'\rangle|^2\int_{-\infty}^{\infty}[1-\cos x]/x^2\, dx$$
$$= 2\pi t\hbar^{-1}|\langle E'\beta''|V|\alpha'\rangle|^2. \quad (38)$$

Luego, la probabilidad de que tenga lugar una transición al estado final en el que las β tengan los valores β'' durante el intervalo de 0 a t es proporcional a t. Existe, por lo tanto, un *coeficiente de probabilidad* definido, o sea, una probabilidad por unidad de tiempo determinada, para el proceso de transición considerado, que tiene el valor

$$2\pi\hbar^{-1}|\langle E'\beta''|V|\alpha'\rangle|^2. \quad (39)$$

Es proporcional al cuadrado del módulo del elemento de matriz de la energía de perturbación asociado con dicha transición.

Si el elemento de matriz $\langle E'\beta''|V|\alpha'\rangle$ es pequeño en comparación con los otros elementos de matriz de V, tendremos que hacer el estudio con la fórmula más precisa (26). Según (35), tenemos

$$\int_0^t \langle\alpha''|V^*(t')|\alpha'''\rangle\, dt' \int_0^{t'} \langle\alpha'''|V^*(t'')|\alpha'\rangle\, dt''$$
$$= \langle\alpha''|V|\alpha'''\rangle\langle\alpha'''|V|\alpha'\rangle \int_0^t e^{i(E''-E''')t'/\hbar}\, dt' \int_0^{t'} e^{i(E'''-E')t''/\hbar}\, dt''$$
$$= \frac{\langle\alpha''|V|\alpha'''\rangle\langle\alpha'''|V|\alpha'\rangle}{i(E'''-E')/\hbar}\int_0^t \{e^{i(E''-E')t'/\hbar} - e^{i(E''-E''')t'/\hbar}\}\, dt'.$$

Si E'' es próximo a E', únicamente el primer término que figura bajo el signo integral en la fórmula anterior da cuenta de una probabilidad de

transición con importancia física, en cambio, el segundo puede ser despreciado. Utilizando este resultado en (26) obtenemos

$$P(\alpha'\alpha'') = 2\left|\langle\alpha''|V|\alpha'\rangle - \sum_{\alpha'''\neq\alpha',\alpha''} \frac{\langle\alpha''|V|\alpha'''\rangle\langle\alpha'''|V|\alpha'\rangle}{E'''-E'}\right|^2 \frac{1-\cos\{(E''-E')t/\hbar\}}{(E''-E')^2}$$

que sustituye a (36). Procediendo igual que antes, obtenemos para la probabilidad de transición por unidad de tiempo a un estado final en el que las β tienen los valores β'' y E tiene un valor próximo a su valor inicial E', el valor

$$\frac{2\pi}{\hbar}\left|\langle E'\beta''|V|\alpha'\rangle - \sum_{\alpha'''\neq\alpha',\alpha''} \frac{\langle E'\beta''|V|\alpha'''\rangle\langle\alpha'''|V|\alpha'\rangle}{E'''-E'}\right|^2. \tag{40}$$

Esta fórmula nos muestra el papel que juegan los estados intermedios distintos del estado inicial y del final en la determinación de un coeficiente de probabilidad.

Para que las aproximaciones utilizadas al deducir (39) y (40) sean válidas, el tiempo t no ha de ser ni muy pequeño ni muy grande. Tiene que ser grande comparado con los períodos de los sistemas atómicos a fin de que el cálculo aproximado de la integral (37) conduzca verdaderamente al valor (38), pero no ha de ser excesivamente grande, pues en tal caso la fórmula general (24) o (26) dejaría de ser válida. De hecho, podríamos hacer la probabilidad (38) mayor que la unidad si tomásemos un tiempo suficientemente grande. El límite superior de t viene determinado por la condición de que la probabilidad (24) o (26), o bien t multiplicado por (39) o (40), sean pequeñas en comparación con la unidad. No hay dificultad para que t satisfaga simultáneamente ambas condiciones siempre que la energía de perturbación V sea suficientemente pequeña.

47. *El efecto Zeeman anómalo*

Uno de los ejemplos más sencillos del método de perturbaciones de § 43 consiste en calcular el cambio de primer orden en los niveles de energía de un átomo producido por un campo magnético uniforme. En § 41 hemos estudiado ya el caso del átomo de hidrógeno, y era tan sencillo que lo hemos podido resolver sin recurrir a la teoría de perturbaciones. El caso de un átomo cualquiera no es mucho más complicado si previamente hacemos una serie de aproximaciones que nos proporcionen un modelo atómico sencillo.

En primer lugar consideremos el átomo en ausencia de campo magnético y busquemos constantes del movimiento o cantidades que sean aproximadamente constantes del movimiento. Ciertamente, el momento angular total del átomo, o sea el vector **j**, es una constante del movimiento. Dicho momento angular puede ser considerado como la suma de dos partes, que son el momento angular orbital total para todos los electrones, que llamaremos **l**, y el momento angular total de spin, que llamaremos **s**. Tenemos, pues, $\mathbf{j} = \mathbf{l} + \mathbf{s}$. Pero el efecto de los momentos magnéticos de spin sobre el movimiento de los electrones es pequeño en comparación con el efecto de las fuerzas de Coulomb, y en primera aproximación puede ser despreciado. En dicha aproximación el momento angular de spin de cada electrón es una constante del movimiento, ya que no hay fuerzas que le obliguen a cambiar su orientación. Así pues, **s**, y, por tanto, también **l** serán constantes del movimiento. Los módulos l, s y j de **l**, **s** y **j** vendrán dados por

$$l + \tfrac{1}{2}\hbar = (l_x^2 + l_y^2 + l_z^2 + \tfrac{1}{4}\hbar^2)^{\frac{1}{2}},$$
$$s + \tfrac{1}{2}\hbar = (s_x^2 + s_y^2 + s_z^2 + \tfrac{1}{4}\hbar^2)^{\frac{1}{2}},$$
$$j + \tfrac{1}{2}\hbar = (j_x^2 + j_y^2 + j_z^2 + \tfrac{1}{4}\hbar^2)^{\frac{1}{2}},$$

según la ecuación (39) de § 36. Todos ellos conmutan entre sí, y según (47) de § 36 vemos que para valores de l y s dados, los posibles valores de j son

$$l + s, \quad l + s - \hbar, \quad \ldots, \quad |l - s|.$$

Consideremos un estado estacionario para el cual l, s y j tengan valores numéricos de acuerdo con este esquema. La energía de dicho estado dependerá de l, pero podría pensarse que despreciando los momentos magnéticos del spin no tendría que depender ni de s ni de la dirección relativa del vector **s** respecto a **l**, es decir, que tampoco habría de depender de j. Sin embargo, veremos en el capítulo IX, que la energía depende considerablemente del módulo s del vector **s**, si bien es independiente de su dirección cuando se desprecia el efecto de los momentos magnéticos del spin debido a ciertos fenómenos que se derivan de la indistinguibilidad de los electrones entre sí. Así pues, existen distintos niveles de energía del sistema para cada valor distinto de l y de s. Esto significa que tanto l como s son funciones de la energía según la definición general de función dada en § 11. pues los valores de l y de s de un estado estacionario quedan fijados cuando se da el valor de la energía.

Ahora podemos tener en cuenta el efecto del momento magnético de spin, considerándolo como una pequeña perturbación según el método de § 43. La energía del sistema no perturbado continuará siendo aproximadamente una constante del movimiento, y en consecuencia, tanto l como s, por ser funciones de la energía, seguirán siendo también aproximadamente

constantes del movimiento. Pero en cambio, las direcciones de los vectores **l** y **s**, como no son funciones de la energía no perturbada, no tienen por qué ser ahora constantes del movimiento ni siquiera en primera aproximación, y pueden estar sometidas a grandes variaciones seculares. Pero como el vector **j** es constante, la única variación posible de **l** y **s** es un movimiento de precesión en torno del vector **j**. Así tenemos un modelo aproximado del átomo en el cual los dos vectores **l** y **s** tienen longitudes constantes y poseen un movimiento de precesión en torno al vector suma de ellos **j**, que es constante. La energía está determinada principalmente por los módulos de **l** y de **s**, y sólo depende ligeramente de la orientación relativa, dada por j. Por tanto, los estados con el mismo l, el mismo s y distinto j pertenecerán a niveles de energía que diferirán muy poco y que constituyen los llamados términos *multipletes*.

Consideremos ahora este modelo atómico como nuestro sistema no perturbado y sometámoslo a un campo magnético uniforme de magnitud $\mathcal{H}$ en la dirección del eje z. La contribución de cada electrón a la energía extra debida al campo magnético vendrá dada por un término

$$e\mathcal{H}/2mc \cdot (m_z + \hbar\,\sigma_z), \tag{41}$$

análogo al último término de la ecuación (89) de § 41. Así pues, en conjunto dicha energía será

$$e\mathcal{H}/2mc \cdot \sum (m_z + \hbar\sigma_z) = e\mathcal{H}/2mc \cdot (l_z + 2s_z) = e\mathcal{H}/2mc \cdot (j_z + s_z). \tag{42}$$

que será la energía de perturbación V. Utilicemos ahora el método de § 43 para determinar los cambios de los niveles de energía producidos por dicha V. El método únicamente será aplicable cuando el campo magnético sea suficientemente pequeño para que V sea pequeño comparado con las diferencias de energía del multiplete.

Nuestro sistema no perturbado es degenerado, dado que la dirección del vector **j** está indeterminada. Por tanto, hemos de tomar aquellos elementos de matriz del representante de V en una representación de Heisenberg del sistema no perturbado, que pertenecen a un mismo nivel de energía respecto a filas y columnas, y hallar los autovalores de la matriz así formada. La forma más cómoda de hacerlo es separar V en dos partes, una que sea constante del movimiento del sistema no perturbado cuyo representante, en consecuencia, sólo tendrá elementos de matriz que correspondan a un mismo nivel de energía respecto a filas y a columnas, y otra cuyo representante no contenga más que elementos de matriz que hagan referencia a dos niveles de energía del sistema no perturbado distintos respecto a filas y a columnas, y que en consecuencia no afectará a la perturbación de primer orden. El término relativo a j_z en (42) es una cons-

tante del movimiento no perturbado, y por tanto pertenece enteramente a la primera parte. Para el término relativo a s_z tenemos

$$s_z(j_x^2 + j_y^2 + j_z^2) = j_z(s_x j_x + s_y j_y + s_z j_z) + (s_z j_x - j_z s_x) j_x + (s_z j_y - j_z s_y) j_y$$

o sea,

$$s_z = \frac{j_z}{j(j+\hbar)} \tfrac{1}{2}[j(j+\hbar) - l(l+\hbar) + s(s+\hbar)] - [\gamma_y j_x - \gamma_x j_y] \frac{1}{j(j+\hbar)} \tag{43}$$

donde

$$\left.\begin{aligned} \gamma_x &= s_z j_y - j_z s_y = s_z l_y - l_z s_y = l_y s_z - l_z s_y, \\ \gamma_y &= j_z s_x - s_z j_x = l_z s_x - s_z l_x = l_z s_x - l_x s_z, \end{aligned}\right\} \tag{44}$$

El primer término de la expresión de s_z es una constante del movimiento no perturbado y, por tanto, pertenece enteramente a la primera parte, mientras que el segundo término pertenece por completo o la segunda parte como vamos a ver a continuación.

En correspondencia con (44), podemos introducir

$$\gamma_z = l_x s_y - l_y s_x.$$

Es fácil comprobar ahora que

$$j_x \gamma_x + j_y \gamma_y + j_z \gamma_z = 0$$

y según (30) de § 35

$$[j_z, \gamma_x] = \gamma_y, \qquad [j_z, \gamma_y] = -\gamma_x, \qquad [j_z, \gamma_z] = 0.$$

Estas relaciones entre j_x, j_y, j_z y γ_x, γ_y, γ_z son de la misma forma que las relaciones entre m_x, m_y, m_z y x, y, z que aparecían en el cálculo de las reglas de selección para los elementos de matriz de z en una representación con k diagonal que hicimos en § 40. Allí obtuvimos el resultado de que todos los elementos de matriz z eran nulos excepto los que hacían referencia a los valores de k que diferían en $\pm\hbar$; así pues, podemos inferir que todos los elementos de matriz de γ_z en una representación con k diagonal, y análogamente los de γ_x y γ_y, son nulos, excepto aquellos que corresponden a valores de j que difieren en $\pm\hbar$. Los coeficientes de γ_x y γ_y del segundo miembro de (43) conmutan con j; y así, el representante del conjunto de dichos términos no contendrá más elementos de matriz que los que correspondan a valores de j que difieren en $\pm\hbar$, y que, por tanto, correspondan a dos niveles de energía del sistema no perturbado distintos.

Luego, si despreciamos la parte de la energía V cuyo representante consta de elementos de matriz que hacen referencia a dos niveles de energía del sistema no perturbado distintos, la energía de perturbación es

$$\frac{e\mathcal{H}}{2mc} j_z \left\{ 1 + \frac{j(j+\hbar) - l(l+\hbar) + s(s+\hbar)}{2j(j+\hbar)} \right\} \tag{45}$$

Los autovalores de este operador dan los cambios de primer orden de los niveles de energía. Podemos hacer que el representante de dicha expresión sea diagonal, eligiendo la representación de tal manera que j_z sea diagonal, y entonces dicha fórmula nos da directamente los cambios de primer orden en los niveles de energía producidos por el campo magnético. Esta expresión se conoce con el nombre de fórmula de Landé.

El resultado (45) es válido únicamente si la energía de perturbación V es pequeña en comparación con las diferencias de energía de un multiplete. Para valores de V mayores es necesario emplear una teoría más complicada. Sin embargo, para campos muy fuertes, cuya energía sea mucho mayor que la correspondiente a las diferencias entre los términos de un multiplete, es posible establecer nuevamente una teoría sencilla. En este caso podemos despreciar la energía de los momentos magnéticos de spin para el átomo en ausencia de campo magnético exterior, con lo que resulta que en el sistema no perturbado de ahora no sólo son constantes los módulos l y s de los vectores $\mathbf{l}$ y $\mathbf{s}$, sino también los propios vectores $\mathbf{l}$ y $\mathbf{s}$. La energía de perturbación V sigue siendo como antes $e\mathcal{H}/2mc \cdot (j_z + s_z)$, pero ahora es una constante del movimiento del sistema no perturbado y sus autovalores nos dan directamente los cambios de los niveles de energía. Dichos autovalores son múltiplos enteros o bien múltiplos impares semienteros de $e\mathcal{H}/2mc$ según que el número de electrones del átomo sea par o impar.

VIII

PROBLEMAS DE COLISIÓN

48. *Observaciones generales*

En este capítulo vamos a estudiar problemas en los que una partícula que proviene del infinito choca o 'entra en colisión' con un sistema atómico, y que tras ser dispersada (scattered) bajo un cierto ángulo continúa nuevamente hacia el infinito. Para abreviar, al sistema atómico que produce la dispersión (scattering) le llamaremos *dispersor* (scatterer). Por tanto, tenemos un sistema constituido por una partícula incidente y un dispersor que interaccionan entre sí, y queremos estudiarlo de acuerdo con las leyes de la mecánica cuántica; deseamos calcular en particular la probabilidad de que la dispersión se realice bajo un ángulo dado. Normalmente se supone que el dispersor tiene masa infinita y que permanece en reposo durante el proceso de dispersión. Este problema fue resuelto por primera vez por Born mediante un método que en lo fundamental es equivalente al que emplearemos en la próxima sección. Hemos de tener en cuenta la posibilidad de que el dispersor considerado como un sistema independiente tenga distintos estados estacionarios, y que si inicialmente estaba en uno de ellos, después de la colisión puede haber quedado en un estado estacionario distinto. Es decir, que la partícula incidente puede inducir transiciones en el dispersor.

El hamiltoniano del sistema total constituido por el dispersor y la partícula incidente no será función explícita del tiempo y, en consecuencia, el sistema total poseerá estados estacionarios representados por soluciones periódicas de la ecuación de ondas de Schrödinger. Pero el significado de estos estados estacionarios requiere un poco de cuidado para ser entendido correctamente. Es evidente que para *cualquier* estado de movimiento del sistema, la partícula pasará casi todo el tiempo en el infinito y, por tanto, el promedio temporal de la posibilidad de encontrar a la partícula en un volumen finito cualquiera será nulo. Ahora bien, para un estado *estacionario* la probabilidad de que la partícula se encuentre en un volumen finito dado, así como cualquier otra observación, ha de ser independiente del tiempo y, por tanto, ha de ser igual al promedio temporal que como

hemos visto es nulo. Por tanto, lo único que tiene interés físicamente son las probabilidades relativas de encontrar la partícula en los diferentes volúmenes finitos, ya que sus valores absolutos son nulos. Dado que la energía inicial de la partícula podía ser cualquiera, la energía total del sistema tendrá un espectro de autovalores continuo. Luego, un ket $|s\rangle$ correspondiente a un estado estacionario de la energía total tendrá longitud infinita. También podemos dar una razón física de ello, y es que si $|s\rangle$ estuviese normalizado y si Q es un observable — función de la posición de la partícula — que vale uno cuando la partícula está en un volumen finito dado y cero en cualquier otro caso, $\langle s|Q|s\rangle$ sería cero, lo que significaría que el valor esperado de Q, o sea, la probabilidad de que la partícula esté en un determinado volumen, es cero. Tal ket no sería útil. Sin embargo, si $|s\rangle$ tiene longitud infinita, $\langle s|Q|s\rangle$ puede ser un número finito, y representará la probabilidad relativa de que la partícula se encuentre en el citado volumen.

Cuando representamos el estado de un sistema por un ket $|x\rangle$ no normalizado para el cual se verifica $\langle x|x\rangle = n$, es conveniente suponer que tenemos n sistemas idénticos que ocupan todos el mismo espacio pero que no interaccionan entre sí, de manera que cada uno de ellos sigue su propio movimiento independientemente de los otros, igual que en el conjunto de Gibbs de § 33. En ese caso podemos interpretar $\langle x|\alpha|x\rangle$ para cualquier observable α, directamente como el valor total de α para los n sistemas. Si aplicamos estas ideas al $|s\rangle$ de antes de longitud infinita correspondiente a un estado estacionario del sistema constituido por el dispersor y la partícula inciden, tendremos que imaginar un conjunto infinito de dichos sistemas que tengan todos ellos el dispersor en el mismo punto del espacio y las partículas distribuidas por todo el espacio de forma continua. El número de partículas en el volumen finito dado será entonces $\langle s|Q|s\rangle$, en donde Q es el observable definido antes, que vale uno si la partícula se encuentra en el volumen dado y cero en cualquier otro caso. Si el ket está representado por una función de onda de Schrödinger relativa a las coordenadas cartesianas de la partícula, entonces el cuadrado del módulo de la función de onda puede ser interpretado directamente como la densidad de partículas en dicha imagen. Sin embargo, debemos recordar que *cada una de dichas partículas tiene su propio dispersor*. Las distintas partículas pueden pertenecer a dispersores que están en estados diferentes. Por tanto, existirá una densidad de partículas distintas para cada estado del dispersor, que será la densidad de partículas que pertenezcan a dispersores que estén en dicho estado. Esto lo tendremos en cuenta considerando que la función de onda es función no sólo de las variables que describen la posición de la partícula, sino también de las que describen el estado del dispersor.

Para hallar coeficientes de dispersión tendremos que estudiar los *estados estacionarios* del sistema total constituido por el dispersor y la partícula.

Por ejemplo, si queremos hallar la probabilidad de que la dispersión se realice según las distintas direcciones cuando el dispersor esté inicialmente en un estado estacionario dado y la partícula tenga inicialmente una velocidad conocida en una dirección también dada, tendremos que estudiar el estado estacionario del sistema total cuya imagen, según el método indicado, represente únicamente partículas pertenecientes a dispersores que estén cada uno en el estado estacionario dado, que a grandes distancias del lugar en que se hallan los dispersores se muevan con la velocidad dada y en la dirección dada, y partículas que pertenezcan posiblemente a dispersores en distintos estados estacionarios, que se muevan *hacia afuera* a partir del punto en que se hallan los dispersores. Esta imagen corresponde muy de cerca al estado de cosas real en una determinación experimental de los coeficientes de dispersión, con la única diferencia de que dicha imagen describe ahora un único sistema real constituido por dispersor y partícula. La distribución de las partículas que se mueven hacia afuera en el infinito nos da directamente, en la imagen, toda la información que se puede obtener experimentalmente acerca de los coeficientes de dispersión. Para calcular prácticamente el estado estacionario descrito en esta imagen, podemos utilizar un método de perturbaciones parecido al de § 43, por ejemplo tomando como sistema no perturbado el sistema constituido por el dispersor y la partícula sin interacción entre ambos.

Estudiando los problemas de colisión, hay todavía otra posibilidad que debemos tener en cuenta, y es que el dispersor puede ser capaz en algunas ocasiones de absorber y reemitir la partícula. Dicha posibilidad se da cuando existen uno o más *estados de absorción* del sistema total, o sea estados que sean *aproximadamente* estacionarios y ligados en el sentido que indicábamos en § 38 (es decir, para los que la probabilidad de que la partícula se encuentre a una distancia del dispersor mayor que r tiende a cero cuando $r \to \infty$). Como los estados de absorción sólo son estacionarios aproximadamente, su propiedad de ser estados ligados solamente será transitoria, y después de un lapso de tiempo suficiente existirá una probabilidad finita de que la partícula escape al infinito. Físicamente esto quiere decir que habrá una probabilidad finita de que la partícula sea emitida espontáneamente. El hecho de que haya sido preciso emplear la palabra 'aproximadamente' para explicar las condiciones requeridas para que puedan ocurrir los fenómenos de emisión y absorción indica que dichas condiciones no se pueden expresar en términos matemáticos exactos. Únicamente podemos dar un significado a dichos fenómenos utilizando un método de perturbaciones. Dichos fenómenos ocurren cuando el sistema no perturbado (constituido por el dispersor y la partícula) tiene estados estacionarios ligados. La introducción de la perturbación destruye las propiedades estacionarias de dichos estados y da lugar a la emisión espontánea y al fenómeno contrario de la absorción.

Contrariamente a lo que ocurría para los coeficientes de dispersión, para calcular probabilidades de absorción y de emisión es necesario considerar *estados no estacionarios*, y, en consecuencia, tendremos que utilizar el método de perturbaciones de § 44. Por lo tanto, para calcular un coeficiente de emisión, tendremos que considerar los estados de absorción no estacionarios que hemos mencionado antes. Además, puesto que toda absorción va seguida siempre de una reemisión, dicho fenómeno no puede ser separado de la dispersión en ningún experimento que no dependa del tiempo explícitamente, correspondiente a un estado estacionario del sistema. Únicamente podemos separar los dos fenómenos en experimentos en que el estado de cosas no se conserva. Por ejemplo, utilizando un haz de partículas incidentes con un comienzo de emisión claramente definido, y así las partículas dispersadas aparecerán inmediatamente después de que las partículas incidentes hayan alcanzado el dispersor, mientras que las que hayan sido absorbidas y después reemitidas aparecerán sólo algún tiempo después. Dicho haz de partículas sería la imagen de un cierto ket de longitud infinita, que puede ser empleado para calcular el coeficiente de absorción.

49. *Coeficiente de dispersión*

Ahora vamos a calcular coeficientes de dispersión, y en primer lugar estudiaremos el caso en que no hay absorción ni emisión, lo que quiere decir que el sistema no perturbado no tiene estados estacionarios ligados. Podemos elegir como sistema no perturbado aquel para el cual no hay interacción entre el dispersor y la partícula. Su hamiltoniano será de la forma

$$E = H_s + W, \tag{1}$$

donde H_s es el hamiltoniano del dispersor considerado él sólo, y W es el de la partícula considerada ella sola. Si despreciamos la mecánica relativista este último vale

$$W = 1/2m \cdot (p_x^2 + p_y^2 + p_z^2). \tag{2}$$

La energía de perturbación V, que suponemos pequeña, será una función de las coordenadas cartesianas de la partícula x, y, z y también quizá de sus momentos p_x, p_y, p_z y de las variables dinámicas que sirven para describir el dispersor.

Puesto que lo que nos interesa son los estados estacionarios del sistema total, emplearemos un método de perturbaciones parecido al de § 43. Pero ahora nuestro sistema no perturbado tendrá necesariamente un intervalo de niveles de energa continuo, pues contiene una partícula libre, y por tanto, tendremos que hacer ciertas modificaciones en nuestro método de

perturbaciones. La cuestión de cómo cambian los niveles de energía debido a la perturbación, que era la base del problema estudiado en § 43, no tiene ahora significado alguno, y el convenio utilizado en § 43 de designar los autovalores de E y de H próximos entre sí con el mismo número de primas es inaplicable. Además, el desdoblamiento de los niveles de energía que tenía lugar en § 43 cuando el sistema no perturbado era degenerado tampoco puede ocurrir ahora, pues cuando el sistema no perturbado sea degenerado, el sistema perturbado deberá tener el mismo grado de degeneración, ya que ha de tener un intervalo de niveles de energía continuo.

Continuaremos utilizando el esquema general de ecuaciones del principio de § 43, que allí correspondía a las ecuaciones (1) a (4); pero ahora tomaremos como estado estacionario del sistema no perturbado que constituye la aproximación de orden cero un estado que pertenezca a un nivel de energía E' precisamente igual al nivel de energía H' del estado estacionario del sistema perturbado. Luego, las a_i introducidas en la segunda de las ecuaciones (3) de § 43 serán ahora todas nulas, y la segunda de las ecuaciones (4) se convertirá en

$$(E' - E)|1\rangle = V|0\rangle. \tag{3}$$

Análogamente, la tercera de las ecuaciones (4) de § 43 será

$$(E' - E)|2\rangle = V|1\rangle. \tag{4}$$

Procedamos a resolver la ecuación (3) y a obtener el coeficiente de dispersión en la aproximación de primer orden. En § 51 será preciso emplear la ecuación (4).

Sea α un conjunto completo de observables que conmutan que sirven para describir el dispersor y que son constantes del movimiento cuando el dispersor se considera solo, siendo, por tanto, útiles para caracterizar los estados estacionarios del dispersor. Ello exige que H_s conmute con los α y que sea una función de ellos. Ahora podemos elegir una representación del sistema total en la que las α y las coordenadas de la partícula x, y, z sean diagonales. En ella H_s será también diagonal. Sea $\langle \mathbf{x}\alpha'|0\rangle$ el representante de $|0\rangle$, y $\langle \mathbf{x}\alpha'|1\rangle$ el de $|1\rangle$; empleamos la variable única $\mathbf{x}$ en lugar de x, y, z y omitimos la prima para abreviar. Asimismo escribiremos d^3x en lugar del producto $dxdydz$. La ecuación (3) escrita en forma de los representantes se convierte, con la ayuda de (1) y (2), en

$$\{E' - H_s(\alpha') + \hbar^2/2m \cdot \nabla^2\}\langle \mathbf{x}\alpha'|1\rangle = \sum_{\alpha''} \int \langle \mathbf{x}\alpha'|V|\mathbf{x}''\alpha''\rangle d^3x''\langle \mathbf{x}''\alpha''|0\rangle. \tag{5}$$

Supongamos que la partícula incidente tiene el momento $\mathbf{p}^0$ y que el estado estacionario inicial del dispersor es α^0. El estado estacionario de nuestro sistema no perturbado será el que corresponde a $\mathbf{p} = \mathbf{p}^0$ y $\alpha = \alpha^0$; luego, su representante será

$$\langle \mathbf{x}\alpha'|0\rangle = \delta_{\alpha'\alpha^0} \; e^{i(\mathbf{p}^0, \mathbf{x})/\hbar}, \tag{6}$$

y así la ecuación (5) se reduce a

$$\{E' - H_s(\alpha') + \hbar^2/2m \cdot \nabla^2\}\langle \mathbf{x}\alpha'|1\rangle = \int \langle \mathbf{x}\alpha'|V|\mathbf{x}^0\alpha^0\rangle \, d^3x^0 \, e^{i(\mathbf{p}^0, \mathbf{x}^0)/\hbar}$$

o sea,

$$(k^2 + \nabla^2)\langle \mathbf{x}\alpha'|1\rangle = F, \tag{7}$$

donde

$$k^2 = 2m\hbar^{-2}\{E' - H_s(\alpha')\} \tag{8}$$

y

$$F = 2m\hbar^{-2} \int \langle \mathbf{x}\alpha'|V|\mathbf{x}^0\alpha^0\rangle \, d^3x^0 e^{i(\mathbf{p}^0, \mathbf{x}^0)/\hbar}, \tag{9}$$

es una función definida de x, y, z y α'. Asimismo resulta

$$E' = H_s(\alpha^0) + \mathbf{p}^2/m. \tag{10}$$

Nuestro problema consiste en hallar una solución $\langle \mathbf{x}\alpha'|1\rangle$ de (7) que para valores de x, y, z correspondientes a puntos alejados del dispersor represente únicamente partículas que se muevan hacia afuera. El cuadrado de su módulo, $|\langle \mathbf{x}\alpha'|1\rangle|^2$, nos dará la densidad de partículas dispersadas pertenecientes a dispersores que se hallen en el estado α' cuando la densidad de partículas incidentes es $|\langle \mathbf{x}\alpha'|0\rangle|^2$, que es igual a uno. Si transformamos la ecuación (7) a coordenadas polares r, θ, ϕ, se convierte en

$$\left\{k^2 + \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2 \operatorname{sen}\theta}\frac{\partial}{\partial\theta}\operatorname{sen}\theta\frac{\partial}{\partial\theta} + \frac{1}{r^2\operatorname{sen}^2\theta}\frac{\partial^2}{\partial\phi^2}\right\}\langle r\theta\phi\alpha'|1\rangle = F. \tag{11}$$

Pero teniendo en cuenta la exigencia física de que la energía de interacción entre el dispersor y la partícula ha de tender a cero cuando la distancia entre ambos tiende a infinito, F tenderá a cero cuando $r \to \infty$. Si ahora despreciamos E en (11), obtenemos la solución aproximada para grandes valores de r

$$\langle r\theta\phi\alpha'|1\rangle = u(\theta\phi\alpha')r^{-1}e^{ikr}, \tag{12}$$

donde u es una función arbitraria de θ, ϕ y α', pues si sustituimos esta expresión en el primer miembro de (11) obtenemos un resultado del orden de r^{-3}. Si no despreciamos F, para valores grandes de r la solución de (11) seguirá siendo de la forma (12), siempre que F tienda a cero suficientemente de prisa cuando $r \to \infty$, pero la función u quedará definida y determinada por la solución para valores de r pequeños.

Para valores α' de α tales que k^2 definido en (8) sea positivo, hemos de elegir el k de (12) igual a la raíz cuadrada positiva de k^2 si queremos que (12) represente únicamente partículas que van hacia afuera, es decir, partículas para las cuales la componente radial del momento, que según § 38 vale $p_r - i\hbar r^{-1}$, o sea, $-i\hbar(\partial/\partial r + r^{-1})$, tenga un valor positivo. Con ello la densidad de partículas dispersadas pertenecientes a dispersores en

el estado α', y que es igual al cuadrado del módulo de (12), disminuye según una ley inversamente proporcional al cuadrado del radio cuando r crece, como necesariamente tiene que ocurrir desde el punto de vista físico, y su distribución angular viene dada por $|u(\theta\phi\alpha')|^2$. Además, la magnitud P' del momento de las partículas dispersadas ha de ser igual a $k\hbar$, pues para r grande el momento es radial; luego, su energía será

$$\frac{P'^2}{2m} = \frac{k^2\hbar^2}{2m} = E' - H_s(\alpha') = H_s(\alpha^0) - H_s(\alpha') + \frac{\mathbf{p}^{02}}{2m},$$

aplicando (8) y (10). Esta energía es precisamente igual a la de la partícula incidente $\mathbf{p}^2/2m$ menos el aumento de energía del dispersador $H_s(\alpha') - H_s(\alpha^0)$, resultado que está de acuerdo con la ley le conservación de la energía. Para valores α' de α tales que k^2 resulte negativo, no hay partículas dispersadas, lo que significa que la energía total inicial es insuficiente para que el dispersor pase al estado α'.

Ahora hemos de calcular $u(\theta\phi\alpha')$ para un conjunto de valores α' de α para los que k^2 sea positivo, y obtener así la distribución angular de las partículas dispersadas pertenecientes a dispersores en el estado α'. Bastará con calcular u en la dirección $\theta = 0$ del polo de las coordenadas polares, pues dicha dirección puede ser elegida arbitrariamente. Haremos uso del teorema de Green, que dice que para dos funciones de la posición A y B cualesquiera, la integral de volumen $\int (A\nabla^2 B - B\nabla^2 A)\, d^3x$ extendida a un volumen cualquiera es igual a la integral de superficie $\int (A\partial B/\partial n - B\partial A/\partial n)\, ds$ extendida a la superficie que limita a dicho volumen, en donde $\partial/\partial n$ significa derivada respecto a la normal a la superficie. Tomamos

$$A = e^{-ikr\cos\theta}, \qquad B = \langle r\theta\phi\alpha'|1\rangle$$

y aplicamos el teorema para una esfera de radio grande con centro en el origen. El integrando de la integral de volumen, según (7) u (11), es

$$\begin{aligned} & e^{-ikr\cos\theta}\,\nabla^2\langle r\theta\phi\alpha'|1\rangle - \langle r\theta\phi\alpha'|1\rangle\,\nabla^2 e^{-ikr\cos\theta} \\ &= e^{-ikr\cos\theta}\,(\nabla^2 + k^2)\langle r\theta\phi\alpha'|1\rangle = e^{-ikr\cos\theta}\,F \end{aligned}$$

mientras que el integrando de la integral de superficie se convierte, mediante (12), en

$$\begin{aligned} & e^{-ikr\cos\theta}\frac{\partial}{\partial r}\langle r\theta\phi\alpha'|1\rangle - \langle r\theta\phi\alpha'|1\rangle\frac{\partial}{\partial r}e^{-ikr\cos\theta} \\ &= e^{-ikr\cos\theta}\,u\left(-\frac{1}{r^2} + \frac{ik}{r}\right)e^{ikr} + i\frac{u}{r}e^{ikr}k\cos\theta\, e^{-ikr\cos\theta} \\ &= iku r^{-1}(1+\cos\theta)e^{ikr(1-\cos\theta)} \end{aligned}$$

despreciando r^{-2}. Luego, resulta

$$\int e^{-ikr\cos\theta} F\, d^3x = \int_0^{2\pi} d\phi \int_0^{\pi} r^2 \operatorname{sen}\theta\, d\theta \cdot ikur^{-1}(1+\cos\theta)e^{ikr(1-\cos\theta)},$$

en donde la integral de volumen del primer miembro está extendida a todo el espacio. Integrando por partes respecto a θ, el segundo miembro se transforma en

$$\int_0^{2\pi} d\phi \left\{ [u(1+\cos\theta)e^{ikr(1-\cos\theta)}]_{\theta=0}^{\theta=\pi} - \int_0^{\pi} e^{ikr(1-\cos\theta)} \frac{\partial}{\partial\theta}[u(1+\cos\theta)]\, d\theta \right\}$$

El segundo término que figura entre los paréntesis { } es del orden de r^{-1}, como puede verse mediante nuevas integraciones por partes y, por lo tanto, puede ser despreciado. Así pues llegamos a

$$\int e^{-ikr\cos\theta} F\, d^3x = -2\int_0^{2\pi} d\phi\, u(0\phi\alpha') = -4\pi u(0\phi\alpha'),$$

que nos da el valor de $u(\theta\phi\alpha')$ en la dirección $\theta = 0$.

Puesto que $P' = k\hbar$, este resultado se puede escribir en la forma

$$u(0\phi\alpha') = -(4\pi)^{-1} \int e^{-1P'r\cos\theta/\hbar} F\, d^3x, \tag{13}$$

Si el vector $\mathbf{p}'$ representa el momento de los electrones dispersados en una cierta dirección (y por tanto, tiene un módulo igual a P'), el valor de u en dicha dirección será

$$u(\theta'\phi'\alpha') = -(4\pi)^{-1} \int e^{-i(\mathbf{p}',\,\mathbf{x})/\hbar} F\, d^3x,$$

según se deduce de (13) eligiendo dicha dirección como polo de las coordenadas polares. Resulta pues, mediante (9),

$$\begin{aligned} u(\theta'\phi'\alpha') &= -(2\pi)^{-1}m\hbar^{-2} \iint e^{-i(\mathbf{p}',\,\mathbf{x})/\hbar}\, d^3x\, \langle \mathbf{x}\alpha'|V|\mathbf{x}^0\alpha^0\rangle\, d^3x^0\, e^{i(\mathbf{p}^0,\,\mathbf{x}^0)/\hbar} \\ &= -2\pi mh\langle \mathbf{p}'\alpha'|V|\mathbf{p}^0\alpha^0\rangle, \end{aligned} \tag{14}$$

después de hacer una transformación y pasar de las coordenadas $\mathbf{x}$ a los momentos $\mathbf{p}$ de la partícula utilizando la función de transformación (54) de § 23. La letra $\mathbf{p}$ es una abreviación de las tres componentes del momento.

Así, la densidad de partículas dispersadas que pertenecen a dispersores en el estado α' será $|u(\theta'\phi'\alpha')|^2/r^2$. Como su velocidad es P'/m, el flujo de partículas por unidad de ángulo sólido y por unidad de tiempo en la dirección del vector $\mathbf{p}'$ será $P'/m \cdot |u(\theta'\phi'\alpha')|^2$. Como hemos visto, la densidad de partículas incidentes era igual a uno y, por tanto, el número de partículas incidentes que atraviesan la unidad de superficie en la unidad

de tiempo es igual a su velocidad P^0/m, siendo P^0 la magnitud de $\mathbf{p}^0$. Luego, la sección eficaz para que una partícula incidente sea dispersada bajo un ángulo sólido unidad en la dirección de $\mathbf{p}'$ y que pertenezca a un dispersor en el estado α' será

$$P'/P^0 \cdot |u(\theta'\phi'\alpha')|^2 = 4\pi^2 m^2 h^2 P'/P^0 \cdot |\langle \mathbf{p}'\alpha'|V|\mathbf{p}^0\alpha^0\rangle|^2. \tag{15}$$

Este es el coeficiente de dispersión relativo a las transiciones $\alpha^0 \to \alpha'$ del dispersor. Dicho coeficiente depende del elemento de matriz $\langle \mathbf{p}'\alpha'|V|\mathbf{p}^0\alpha^0\rangle$ de la energía de perturbación V cuya columna $\mathbf{p}^0\alpha^0$ y cuya fila $\mathbf{p}'\alpha'$ se refieren respectivamente a los estados inicial y final del sistema no perturbado entre los que tiene lugar la transición. Así pues, el resultado (15) es análogo al resultado (24) de § 44, si bien los coeficientes numéricos son distintos en ambos casos en correspondencia con la distinta naturaleza de los dos procesos de transición.

50. *Solución en la representación de momentos*

El valor del coeficiente de dispersión dado por (15) sólo hace referencia a la representación en que $\mathbf{p}$ es diagonal. Así pues, es de esperar que podríamos llegar a él de un modo más directo utilizando desde el principio la representación $\mathbf{p}$ en lugar de utilizar la representación $\mathbf{x}$ y transformar al final el resultado expresándolo en la representación $\mathbf{p}$, tal como hemos hecho en § 49. A primera vista no parece que esto constituya una gran ventaja, pues si bien el método de la representación $\mathbf{x}$ no resulta tan directo, ello queda compensado por ser más inmediatamente aplicable, ya que el cuadrado del módulo del representante de un estado en la representación $\mathbf{x}$ nos da la densidad del flujo de partículas en el proceso de dispersión. Sin embargo, el método de la representación $\mathbf{x}$ tiene otros inconvenientes más serios. Una de las principales aplicaciones de la teoría de colisiones corresponde al caso en que las partículas incidentes son fotones. Pero un fotón no es una partícula simple, pues tiene una polarización. Es evidente, a partir de la teoría electromagnética clásica, que un fotón con un momento definido, es decir, que se mueve en una cierta dirección con una frecuencia definida, puede tener un estado de polarización definido (rectilíneo, circular, etc.), mientras que un fotón con una posición definida, cosa que en nuestra imagen quiere decir una alteración electromagnética confinada a un volumen muy pequeño, no puede tener un estado de polarización definido. Estos hechos vienen a decir que el observable polarización de un fotón conmuta con los momentos pero no con la posición. Esto se traduce en que el método de la representación $\mathbf{p}$ es directamente aplicable a fotones si introducimos la variable de polarización en los representantes y la tratamos del mismo modo que tratábamos a

los α que describen el dispersor, mientras que el método de la representación $\mathbf{x}$ no se puede aplicar. Además, cuando se consideran fotones es necesario tener en cuenta la mecánica relativista. En el método de la representación $\mathbf{p}$ es fácil hacerlo, pero en el de la representación $\mathbf{x}$ no resulta tan fácil.

La ecuación (3) sigue siendo válida en mecánica relativista, pero en ella W viene dada por

$$W^2/c^2 = m^2c^2 + P^2 = m^2c^2 + p_x^2 + p_y^2 + p_z^2 \tag{16}$$

en lugar de por (2). Escrita en forma de representantes $\mathbf{p}$, la ecuación (3) es

$$\{E' - H_s(\alpha') - W\}\langle \mathbf{p}\alpha'|1\rangle = \langle \mathbf{p}\alpha'|V|0\rangle,$$

en donde para abreviar ponemos $\mathbf{p}$ en lugar de $\mathbf{p}'$ y sobreentendemos que W es una función determinada de p_x, p_y y p_z dada por (16). Esto se puede escribir

$$(W' - W)\langle \mathbf{p}\alpha'|1\rangle = \langle \mathbf{p}\alpha'|V|0\rangle, \tag{17}$$

donde

$$W' = E' - H_s(\alpha') \tag{18}$$

es la energía de la partícula dispersada perteneciente a un dispersor en el estado α', que corresponde a la ley de conservación de la energía. El representante $\mathbf{x}$ del ket $|0\rangle$ es (6), y el del ket básico $|\mathbf{p}^0\alpha^0\rangle$ es

$$\langle \mathbf{x}\alpha'|\mathbf{p}^0\alpha^0\rangle = \delta_{\alpha'\alpha^0}\ \langle \mathbf{x}|\mathbf{p}^0\rangle = \delta_{\alpha'\alpha^0}\ h^{-\frac{3}{2}}e^{i(\mathbf{p}^0,\ \mathbf{x})/\hbar},$$

según se deduce de la función de transformación (54) de § 23. Luego,

$$|0\rangle = h^{\frac{3}{2}}|\mathbf{p}^0\alpha^0\rangle, \tag{19}$$

y así la ecuación (17) se puede escribir

$$(W' - W)\langle \mathbf{p}\alpha'|1\rangle = h^{\frac{3}{2}}\langle \mathbf{p}\alpha'|V|\mathbf{p}^0\alpha^0\rangle. \tag{20}$$

Ahora hagamos una transformación de coordenadas y pasemos de las componentes cartesianas p_x, p_y y p_z de $\mathbf{p}$ a sus componentes polares P, ω y χ dadas por

$$p_x = P\cos\omega, \qquad p_y = P\,\mathrm{sen}\,\omega\cos\chi, \qquad p_z = P\,\mathrm{sen}\,\omega\,\mathrm{sen}\,\chi.$$

Si en la nueva representación tomamos la función de peso $P^2\,\mathrm{sen}\,\omega$, entonces el peso asignado a cualquier volumen del espacio $\mathbf{p}$ será el mismo que en la representación $\mathbf{p}$ de antes y, en consecuencia, la transformación se reducirá a cambiar los símbolos de las filas y columnas de las matrices sin alterar en nada dichos elementos. Por tanto, en la nueva representación, (20) será

$$(W' - W)\langle P\omega\chi\alpha'|1\rangle = h^{\frac{3}{2}}\langle P\omega\chi\alpha'|V|P^0\omega^0\chi^0\alpha^0\rangle, \tag{21}$$

donde W ahora únicamente será función de P.

El coeficiente $W' - W$ de $\langle P\omega\chi\alpha'|1\rangle$ aquí es simplemente un factor multiplicativo y no un operador de derivación como ocurría en la representación **x**. Así pues, podemos dividir por él y obtener una expresión explícita de $\langle P\omega\chi\alpha'|1\rangle$. Pero si a α' corresponde un W' definido por (18) mayor que mc^2, dicho factor tendrá el valor cero en un cierto punto del dominio de la variable P, que será el punto $P = P'$, dado en función de W' por (16). Por tanto, la función $\langle P\omega\chi\alpha'|1\rangle$ tendrá una singularidad en dicho punto. Dicha singularidad nos indica que el número de partículas que se mueven a grandes distancias de los dispersores con energías infinitamente próximas a W' representado por $\langle P\omega\chi\alpha'|1\rangle$ es infinito y, por tanto, es precisamente dicha singularidad lo que nos interesa para hallar la distribución angular de partículas en el infinito.

El resultado de dividir (21) por el factor $W' - W$, según (13) de § 15, es

$$\langle P\omega\chi\alpha'|1\rangle = h^{\frac{3}{2}}\langle P\omega\chi\alpha'|V|P^0\omega^0\chi^0\alpha^0\rangle/(W' - W) + \lambda(\omega\chi\alpha')\delta(W' - W), \tag{22}$$

donde λ es una función arbitraria de ω, χ y α'. Para dar significado al primer término del segundo miembro de (22), haremos el convenio de que su integral respecto a P extendida a un dominio que contenga el valor P' es igual al límite cuando $\varepsilon \to 0$ de la integral extendida a dicho dominio del que se ha suprimido el pequeño intervalo $(P' - \varepsilon, P' + \varepsilon)$. Esto es suficiente para precisar el significado de (22), pues únicamente nos interesan integrales de los representantes de los estados cuando la representación contiene un conjunto continuo de filas y columnas. Vemos que la ecuación (21) no es suficiente para determinar completamente el representante $\langle P\omega\chi\alpha'|1\rangle$ debido a la función arbitraria λ que aparece en (22). Hemos de elegir λ de forma que $\langle P\omega\chi\alpha'|1\rangle$ represente únicamente partículas que se mueven hacia afuera, pues queremos que las únicas partículas incidentes correspondan al ket $|0\rangle$.

Consideremos en primer lugar el caso en que el representante $\langle P\omega\chi|\rangle$ de un estado de la partícula satisfaga una ecuación general del tipo

$$(W' - W)\langle P\omega\chi|\rangle = f(P\omega\chi), \tag{23}$$

con una $f(P\omega\chi)$ función arbitraria de P, ω y χ, y siendo W' un número mayor que mc^2. $\langle P\omega\chi|\rangle$ será entonces de la forma

$$\langle P\omega\chi|\rangle = f(P\omega\chi)/(W' - W) + \lambda(\omega\chi)\delta(W' - W), \tag{24}$$

y determinamos λ a fin de que $\langle P\omega\chi|\rangle$ represente únicamente partículas que se mueven hacia afuera. Podemos hacerlo expresando $\langle P\omega\chi|\rangle$ en la representación **x** o mejor a la representación $(r\theta\phi)$, y comparándolo con (12) para valores grandes de r. La función de transformación es

$$\langle r\theta\phi|P\omega\chi\rangle = h^{-\frac{3}{2}}e^{i(\mathbf{p},\,\mathbf{x})/\hbar} = h^{-\frac{3}{2}}e^{iPr[\cos\omega\cos\theta + \operatorname{sen}\omega\operatorname{sen}\theta\cos(\chi-\phi)]/\hbar}.$$

Para la dirección $\theta = 0$ resulta

$$\langle r0\phi|\rangle = h^{-\frac{3}{2}} \int_0^\infty P^2\, dP \int_0^{2\pi} d\chi \int_0^\pi \operatorname{sen}\omega\, d\omega\, e^{iPr\cos\omega/\hbar} \langle P\omega\chi|\rangle$$

$$= h^{-\frac{3}{2}} \int_0^\infty P^2\, dP \int_0^{2\pi} d\chi \left\{ -\left[\frac{e^{iPr\cos\omega/\hbar}}{iPr/\hbar} \langle P\omega\chi|\rangle \right]_{\omega=0}^{\omega=\pi} + \right.$$

$$\left. + \int_0^\pi d\omega\, \frac{e^{iPr\cos\omega/\hbar}}{iPh/\hbar} \frac{\partial}{\partial\omega} \langle P\omega\chi|\rangle \right\}$$

Como puede verse integrando nuevamente por partes respecto a ω, el segundo término dentro del paréntesis {} es de orden r^{-2} y, por tanto, puede ser despreciado. Llegamos pues a

$$\langle r0\phi|\rangle = ih^{-\frac{1}{2}}(2\pi r)^{-1} \int_0^\infty P\, dP \int_0^{2\pi} d\chi \{e^{-iPr/\hbar}\langle P\pi\chi|\rangle - e^{iPr/\hbar}\langle P0\chi|\rangle\}$$

$$= ih^{-\frac{1}{2}}r^{-1} \int_0^\infty P\, dP\{e^{-iPr/\hbar}\langle P\pi\chi|\rangle - e^{iPr/\hbar}\langle P0\chi|\rangle\}. \tag{25}$$

Si sustituimos $\langle P\omega\chi|\rangle$ por su valor (24), el primer término del integrando de (25) da

$$ih^{-\frac{1}{2}}r^{-1} \int_0^\infty P\, dP\, e^{-iPr/\hbar}\{f(P\pi\chi)/(W' - W) + \lambda(\pi\chi)\delta(W' - W)\}. \tag{26}$$

El término que lleva consigo la $\delta(W' - W)$ se puede integrar inmediatamente utilizando la relación $P\, dP = W dW/c^2$ que es consecuencia de (16), y da

$$ih^{-\frac{1}{2}}c^{-2}r^{-1} \int_{mc^2}^\infty W\, dW\, e^{-iPr/\hbar}\lambda(\pi\chi)\delta(W' - W)$$

$$= ih^{-\frac{1}{2}}c^{-2}r^{-1}W'\lambda(\pi\chi)e^{-iP'r/\hbar} \tag{27}$$

Para integrar el otro término de (26) emplearemos la fórmula

$$\int_0^\infty g(P)\frac{e^{-iPr/\hbar}}{P' - P}\, dP = g(P') \int_0^\infty \frac{e^{-iPr/\hbar}}{P' - P}\, dP, \tag{28}$$

que es válida, despreciando los términos de orden r^{-1}, para toda función $g(P)$ continua. En efecto, $\int_0^\infty K(P)e^{-iPr/\hbar}\, dP$ es de orden r^{-1} para cualquier función $K(P)$ continua, y la diferencia

$$g(P)/(P' - P) - g(P')/(P' - P)$$

es continua. El segundo miembro de (28) calculado despreciando los tér-

minos de orden r^{-1} y dejando fuera del dominio de integración el pequeño intervalo $(P' - \varepsilon,\ P' + \varepsilon)$, da

$$g(P')\int_{-\infty}^{\infty}\frac{e^{-iPr/\hbar}}{P'-P}\,dP = g(P')e^{-iP'r/\hbar}\int_{-\infty}^{\infty}\frac{e^{i(P'-P)r/\hbar}}{P'-P}\,dP$$

$$= ig(P')e^{-iP'r/\hbar}\int_{-\infty}^{\infty}\frac{\operatorname{sen}(P'-P)r/\hbar}{P'-P}\,dP = i\pi g(P')e^{-iP'r/\hbar}. \qquad (29)$$

En nuestro caso, $g(P)$ es

$$g(P) = ih^{-\frac{1}{2}}r^{-1}P\,f(P\pi\chi)(P'-P)/(W'-W).$$

que para $P = P'$ toma el valor límite

$$g(P') = ih^{-\frac{1}{2}}r^{-1}P'f(P'\pi\chi)W'/P'c^2 = ih^{-\frac{1}{2}}c^{-2}r^{-1}W'f(P'\pi\chi).$$

Sustituyendo esto en (29) y sumándole la expresión (27) obtenemos el siguiente valor de la integral (26)

$$h^{-\frac{1}{2}}c^{-2}r^{-1}W'\{-\pi f(P'\pi\chi) + i\lambda(\pi\chi)\}e^{-iP'r/\hbar}. \qquad (30)$$

Análogamente, el segundo término del integrando de (25) vale

$$h^{-\frac{1}{2}}c^{-2}r^{-1}W'\{-\pi f(P'0\chi) - i\lambda(0\chi)\}e^{iP'r/\hbar}. \qquad (31)$$

La suma de estas dos expresiones es el valor de $\langle r0\phi|\rangle$ para r grande.

$\langle r0\phi|\rangle$ debe representar únicamente partículas que se mueven hacia afuera, luego, ha de ser un múltiplo de $e^{iP'r/\hbar}$. Por tanto, (30) ha de ser nulo, de donde

$$\lambda(\pi\chi) = -\,i\pi f(P'\pi\chi). \qquad (32)$$

Resulta pues, que la condición de que $\langle r0\phi|\rangle$ represente únicamente partículas que vayan hacia afuera en la dirección $\theta = 0$ determina el valor de λ en la dirección opuesta $\theta = \pi$. Pero como la dirección $\theta = 0$ u $\omega = 0$ del polo de las coordenadas polares no es privilegiada, podemos generalizar (32) y escribir

$$\lambda(\omega\chi) = -\,i\pi f(P'\omega\chi), \qquad (33)$$

que nos da el valor de λ en una dirección cualquiera. Dicho valor sustituido en (24) nos da un resultado que se puede escribir en la forma

$$\langle P\omega\chi|\rangle = f(P\omega\chi)\{1/(W'-W) - i\pi\,\delta(W'-W)\}, \qquad (34)$$

pues en el coeficiente de un término que contiene $\delta(W'-W)$ como factor se puede sustituir P' por P sin alterar su valor. Luego, *la condición para que* $\langle P\omega\chi|\rangle$ *represente únicamente partículas que van hacia afuera es que contenga el factor*

$$\{1/(W'-W) - i\pi\,\delta(W'-W)\}. \qquad (35)$$

Observemos que este factor es de la forma del segundo miembro de la ecuación (15) de § 15.

Teniendo en cuenta el valor de λ dado por (33) la expresión (30) es nula, y en consecuencia, el valor de $\langle r0\phi|\rangle$ para r grande viene dado por (31), es decir,

$$\langle r0\phi|\rangle = -2\pi h^{-\frac{1}{2}}c^{-2}r^{-1}W'f(P'0\chi)e^{iP'r/\hbar}.$$

Esto puede generalizarse y resulta

$$\langle r\theta\phi|\rangle = -2\pi h^{-\frac{1}{2}}c^{-2}r^{-1}W'f(P'\omega\chi)e^{iP'r/\hbar},$$

que da el valor de $\langle r\theta\phi|\rangle$ para cualquier dirección θ, ϕ en función de $f(P'\omega\chi)$ en esa misma dirección que designamos por ω, χ. Esta ecuación es de la forma (12) con

$$u(\theta\phi) = -2\pi h^{-\frac{1}{2}}c^{-2}W'f(P'\omega\chi)$$

y por tanto, representa una distribución de partículas de momento P' y de densidad

$$\frac{c^2P'}{W'}|u|^2 = \frac{4\pi^2W'P'}{hc^2}|f(P'\omega\chi)|^2 \tag{36}$$

por unidad de ángulo sólido y de tiempo, que se mueven hacia afuera, que será la distribución representada por $\langle P\omega\chi|\rangle$ de (34).

De esta conclusión general podemos inferir que, siempre que tengamos un representante $\langle P\omega\chi|\rangle$ que represente únicamente partículas que se mueven hacia afuera y que verifica una ecuación del tipo (23), el número de partículas por unidad de ángulo sólido y de tiempo viene dado por (36). Si dicho $\langle P\omega\chi|\rangle$ se refiere a un problema en el que el número de partículas incidentes por unidad de volumen es la unidad, corresponderá a un coeficiente de dispersión cuyo valor es

$$\frac{4\pi^2W^0W'P'}{hc^4P^0}|f(P'\omega\chi)|^2. \tag{37}$$

Sólo depende del valor de la función $f(P'\omega\chi)$ en el punto $P = P'$.

Aplicando esta teoría general a las ecuaciones (21) y (22) tenemos

$$f(P\omega\chi) = h^{\frac{3}{2}}\langle P\omega\chi\alpha'|V|P^0\omega^0\chi^0\alpha^0\rangle.$$

Luego, según (37), el coeficiente de dispersión es

$$4\pi^2h^2W^0W'P'/c^4P^0 \cdot |\langle P'\omega\chi\alpha'|V|P^0\omega^0\chi^0\alpha^0\rangle|^2. \tag{38}$$

Si despreciamos la relatividad y hacemos $W^0W'/c^4 = m^2$, este resultado se reduce a la expresión (15), obtenida en la sección anterior con ayuda del teorema de Green.

51. *Dispersión con absorción y reemisión*

Vamos a hallar ahora el coeficiente de dispersión cuando la partícula puede ser absorbida, es decir, cuando el sistema no perturbado dispersor-partícula tiene estados estacionarios ligados en los cuales la partícula está absorbida. Veremos que la existencia de estos estados ligados del sistema no perturbado tiene un efecto considerable en la dispersión del sistema perturbado, efecto que depende mucho de la energía de la partícula incidente, y que da lugar en óptica al fenómeno de la 'dispersión' † cuando la partícula incidente es un fotón.

Utilizaremos una representación cuyos kets básicos correspondan a los estados estacionarios del sistema no perturbado, como en la representación **p** de la sección anterior. Elegiremos como estados estacionarios del sistema no perturbado los estados $(\mathbf{p}'\alpha')$ en que la partícula tiene un momento $\mathbf{p}'$ definido y el dispersor esté en un estado α' definido, y además los estados ligados k que constituyen un conjunto discreto separado, y supondremos que dichos estados son todos independientes y ortogonales entre sí. Esta hipótesis no es adecuada cuando la partícula es un electrón o un núcleo atómico, pues en ambos casos la partícula sigue teniendo alguna existencia en el estado de absorción k y, por tanto, es de esperar que podremos expresar $|k\rangle$ en función de autokets $|\mathbf{x}'\alpha'\rangle$ de x, y, z y de los α, y en consecuencia, también en función de los $|\mathbf{p}'\alpha'\rangle$. En cambio, si las partículas incidentes son fotones, en los estados de absorción no tendrán existencia alguna y, por tanto, dichos estados serán forzosamente independientes y ortogonales a los estados $(\mathbf{p}'\alpha')$ en que la partícula existe. Luego, la hipótesis es aplicable en este caso de gran interés práctico.

Como se trata de un problema de dispersión, hemos de considerar estados *estacionarios* del sistema total. Pero aquí tendremos que trabajar con la aproximación de segundo orden, pues la ecuación de primer orden (3) no es suficiente, y tendremos que emplear también la (4). La ecuación (3) escrita en función de los representantes de la representación elegida es

$$\left.\begin{aligned}(W'-W)\langle\mathbf{p}\alpha'|1\rangle &= \langle\mathbf{p}\alpha'|V|0\rangle,\\ (E'-E_k)\langle k|1\rangle &= \langle k|V|0\rangle,\end{aligned}\right\}\qquad(39)$$

donde W' es la función de E' y de α' dada por (18), y E_k es la energía del estado estacionario k del sistema no perturbado. Análogamente, la ecuación (4) será

$$\left.\begin{aligned}(W'-W)\langle\mathbf{p}\alpha'|2\rangle &= \langle\mathbf{p}\alpha'|V|1\rangle,\\ (E'-E_k)\langle k|2\rangle &= \langle k|V|1\rangle.\end{aligned}\right\}\qquad(40)$$

† *N. de T.:* Hemos traducido en todo el texto "scattering" por dispersión, pues no existe ningún término latino que corresponda exactamente al mismo concepto. En realidad "dispersión" es un caso particular de "scattering", y en el texto inglés se emplea exclusivamente en esta ocasión.

Desarrollando el segundo miembro mediante multiplicación de matrices resulta

$$\left.\begin{aligned}(W'-W)\langle \mathbf{p}\alpha'|2\rangle &= \sum_{\alpha''}\int \langle \mathbf{p}\alpha'|V|\mathbf{p}''\alpha''\rangle\, d^3p''\, \langle \mathbf{p}''\alpha''|1\rangle + \sum_{k''} \langle \mathbf{p}\alpha'|V|k''\rangle\langle k''|1\rangle,\\ (E'-E_k)\langle k|2\rangle &= \sum_{\alpha''}\int \langle k|V|\mathbf{p}''\alpha''\rangle\, d^3p''\, \langle \mathbf{p}''\alpha''|1\rangle + \sum_{k''} \langle k|V|k''\rangle\langle k''|1\rangle.\end{aligned}\right\} \quad (41)$$

El ket $|0\rangle$ vendrá dado, como antes, por (19), de donde (39) se podrá escribir

$$(W'-W)\langle \mathbf{p}\alpha'|1\rangle = h^{\frac{3}{2}}\langle \mathbf{p}\alpha'|V|\mathbf{p}^0\alpha^0\rangle, \quad (42)$$

$$(E'-E_k)\langle k|1\rangle = h^{\frac{3}{2}}\langle k|V|\mathbf{p}^0\alpha^0\rangle. \quad (43)$$

Podemos suponer que los elementos de matriz $\langle k'|V|k''\rangle$ de V son nulos, puesto que dichos elementos no son esenciales para el fenómeno que estamos investigando, y de no ser nulos no significaría más que que no habríamos elegido adecuadamente los estados de absorción k. Supondremos además que si se consideran los elementos de matriz $\langle k'|V|\mathbf{p}''\alpha''\rangle$, $\langle \mathbf{p}'\alpha'|V|k''\rangle$ de primer orden de magnitud, los elementos $\langle \mathbf{p}'\alpha'|V|\mathbf{p}''\alpha''\rangle$ son de segundo orden. Esta hipótesis la justificaremos en el caso de fotones en § 64. Con ello, según (43) y (42), resulta $\langle k|1\rangle$ de primer orden siempre que E' no tenga un valor próximo a alguno de los niveles de energía discretos E_k, y en cambio $\langle \mathbf{p}\alpha'|1\rangle$ será de segundo orden. Por tanto, según la primera ecuación (41), $\langle \mathbf{p}\alpha'|2\rangle$ en segunda aproximación será

$$(W'-W)\langle \mathbf{p}\alpha'|2\rangle = h^{\frac{3}{2}} \sum_{k''} \langle \mathbf{p}\alpha'|V|k''\rangle\langle k''|V|\mathbf{p}^0\alpha^0\rangle/(E'-E_{k''}).$$

Luego, la corrección total hasta el segundo orden de la función de onda, es decir $\langle \mathbf{p}\alpha'|1\rangle$ más $\langle \mathbf{p}\alpha'|2\rangle$, verificará

$$\begin{aligned}(W'-W)\{\langle \mathbf{p}\alpha'|1\rangle + \langle \mathbf{p}\alpha'|2\rangle\}\\ = h^{\frac{3}{2}}\{\langle \mathbf{p}\alpha'|V|\mathbf{p}^0\alpha^0\rangle + \sum_k \langle \mathbf{p}\alpha'|V|k\rangle\langle k|V|\mathbf{p}^0\alpha^0\rangle/(E'-E_k)\}.\end{aligned}$$

Esta ecuación es del tipo (23) si α' es tal que $W' > mc^2$, lo que es equivalente a exigir que el estado final α' sea compatible con la ley de conservación de la energía. Luego, según el resultado general (37), podemos deducir que el coeficiente de dispersión es

$$\frac{4\pi^2h^2W^0W'P'}{c^4P^0}\left|\langle \mathbf{p}'\alpha'|V|\mathbf{p}^0\alpha^0\rangle + \sum_k \frac{\langle \mathbf{p}'\alpha'|V|k\rangle\langle k|V|\mathbf{p}^0\alpha^0\rangle}{E'-E_k}\right|^2. \quad (44)$$

En este caso la dispersión puede considerarse compuesta por dos partes, una que proviene del elemento de matriz $\langle \mathbf{p}'\alpha'|V|\mathbf{p}^0\alpha^0\rangle$ de la energía de perturbación, y otra que se deriva de los elementos de matriz $\langle \mathbf{p}'\alpha'|V|k\rangle$ y

$\langle k|V|\mathbf{p}^0\alpha^0\rangle$. La primera parte, que es igual al valor obtenido previamente en (38) podemos llamarla dispersión directa. La segunda puede considerarse producida por una absorción de la partícula incidente pasando el sistema a un estado k seguida inmediatamente de una reemisión en una dirección distinta, fenómeno análogo a las transiciones a través de estados intermedios considerados en § 44. El hecho de que hayamos tenido que sumar los dos términos antes de tomar el cuadrado del módulo nos dice que hay interferencia entre los dos tipos de dispersión. No existe ningún procedimiento experimental para separar las dos clases de dispersión, y la distinción entre ambas es de índole puramente matemática.

52. *Dispersión de resonancia*

Supongamos ahora que variamos continuamente la energía de la partícula incidente dejando fijo el estado inicial del dispersor α^0, y así la energía total E' o H' variará continuamente. La fórmula (44) nos dice que cuando E' se aproxima mucho a uno de los niveles de energía discretos E_k, la dispersión se hace muy grande. De hecho, según la fórmula (44), la dispersión debiera ser infinita cuando E' fuera exactamente igual a un E_k. Por supuesto, un coeficiente de dispersión infinito es imposible desde el punto de vista físico, lo que nos induce a pensar que la aproximación empleada en la deducción de (44) deja de ser válida cuando E' está próximo a algún E_k. Para estudiar la dispersión en este caso hemos de recurrir a la ecuación exacta (2) de § 43 con E' en lugar de H',

$$(E' - E)|H'\rangle = V|H'\rangle,$$

y emplear un método distinto para aproximar la solución. Esta ecuación exacta escrita como (41) en forma de representantes es

$$\left.\begin{aligned}
&(W' - W)\langle \mathbf{p}\alpha'|H'\rangle \\
&= \sum_{\alpha''} \int \langle \mathbf{p}\alpha'|V|\mathbf{p}''\alpha''\rangle\, d^3p''\, \langle \mathbf{p}''\alpha''|H'\rangle + \sum_{k''} \langle \mathbf{p}\alpha'|V|k''\rangle\langle k''|H'\rangle, \\
&(E' - E_k)\langle k|H'\rangle \\
&= \sum_{\alpha''} \int \langle k|V|\mathbf{p}''\alpha''\rangle\, d^3p''\, \langle \mathbf{p}''\alpha''|H'\rangle + \sum_{k''} \langle k|V|k''\rangle\langle k''|H'\rangle.
\end{aligned}\right\} \quad (45)$$

Tomemos un E_k particular y consideremos el caso en que E' sea próximo a él. El término grande del coeficiente de dispersión (44) proviene de los elementos de matriz de V de la fila k o de la columna k, es decir, de los elementos del tipo $\langle k|V|\mathbf{p}\alpha'\rangle$ o $\langle \mathbf{p}\alpha'|V|k\rangle$. La dispersión debida a los otros elementos de matriz de V es de un orden de magnitud inferior. Ello nos sugiere que la aproximación que debemos hacer en la ecuación exacta (45)

es despreciar todos los elementos de matriz de V salvo los importantes, es decir, salvo los del tipo $\langle \mathbf{p}\alpha'|V|k\rangle$ o $\langle k|V|\mathbf{p}\alpha'\rangle$, donde α' es un estado final del dispersor cuya energía no supera el valor máximo compatible con la ley de conservación de la energía. En dicha aproximación, las ecuaciones se reducen a

$$\left.\begin{aligned} (W'-W)\langle \mathbf{p}\alpha'|H'\rangle &= \langle \mathbf{p}\alpha'|V|k\rangle\langle k|H'\rangle, \\ (E'-E_k)\langle k|H'\rangle &= \sum_{\alpha'} \int \langle k|V|\mathbf{p}\alpha'\rangle\, d^3p\, \langle \mathbf{p}\alpha'|H'\rangle, \end{aligned}\right\} \quad (46)$$

en donde la sumación respecto de α' se refiere a aquellos valores de α' para los cuales W' dado por (18) es mayor que mc^2. Estas ecuaciones son suficientemente sencillas para poder ser resueltas exactamente sin tener que hacer nuevas aproximaciones.

A partir de la primera ecuación (46), por división obtenemos

$$\langle \mathbf{p}\alpha'|H'\rangle = \langle \mathbf{p}\alpha'|V|k\rangle\langle k|H'\rangle/(W'-W) + \lambda\,\delta(W'-W). \quad (47)$$

Hemos de elegir λ — que puede ser una función cualquiera de los momentos $\mathbf{p}$ y de α' — de modo que (47) represente las partículas incidentes correspondientes a $|0\rangle$ o $h^{\frac{3}{2}}|\mathbf{p}^0\alpha^0\rangle$ y partículas que vayan todas hacia afuera. [El representante de $h^{\frac{3}{2}}|\mathbf{p}^0\alpha^0\rangle$ ahora es de la forma $\lambda\,\delta(W'-W)$, ya que las condiciones $\alpha' = \alpha^0$ y $\mathbf{p} = \mathbf{p}^0$ para que no sea nulo nos llevan a $W' = E' - H_s(\alpha') = E' - H_s(\alpha^0) = W^0 = W$.] Luego, (47) ha de ser

$$\begin{aligned} \langle \mathbf{p}\alpha'|H'\rangle = {} & h^{\frac{3}{2}}\langle \mathbf{p}\alpha'|\mathbf{p}^0\alpha^0\rangle + \\ & + \langle \mathbf{p}\alpha'|V|k\rangle\langle k|H'\rangle\{1/(W'-W) - i\pi\,\delta(W'-W)\}, \end{aligned} \quad (48)$$

y según la fórmula general (37), el coeficiente de dispersión será

$$4\pi^2 W^0 W' P'/hc^4 P^0 \cdot |\langle \mathbf{p}'\alpha'|V|k\rangle|^2 |\langle k|H'\rangle|^2. \quad (49)$$

Nos queda por calcular el valor de $\langle k|H'\rangle$. Para ello podemos sustituir $\langle \mathbf{p}\alpha'|H'\rangle$ en la segunda ecuación (46) por su valor (48). Así resulta

$$\begin{aligned} (E'-E_k)\langle k|H'\rangle &= h^{\frac{3}{2}}\langle k|V|\mathbf{p}^0\alpha^0\rangle + \\ &+ \langle k|H'\rangle \sum_{\alpha'} \int |\langle k|V|\mathbf{p}\alpha'\rangle|^2\{1/(W'-W) - i\pi\,\delta(W'-W)\}\, d^3p \\ &= h^{\frac{3}{2}}\langle k|V|\mathbf{p}^0\alpha^0\rangle + \langle k|H'\rangle(a - ib), \end{aligned}$$

donde

$$a = \sum_{\alpha'} \int |\langle k|V|\mathbf{p}\alpha'\rangle|^2\, d^3p/(W'-W) \quad (50)$$

y

$$b = \pi \sum_{\alpha'} \int |\langle k|V|\mathbf{p}\alpha'\rangle|^2 \, \delta(W' - W) \, d^3p$$

$$= \pi \sum_{\alpha'} \iiint |\langle k|V|P\omega\chi\alpha'\rangle|^2 \, \delta(W' - W) P^2 \, dP \operatorname{sen} \omega \, d\omega d\chi$$

$$= \pi \sum_{\alpha'} P'W'c^{-2} \iint |\langle k|V|P'\omega\chi\alpha'\rangle|^2 \operatorname{sen} \omega \, d\omega d\chi. \tag{51}$$

Luego,

$$\langle k|H'\rangle = h^3\langle k|V|\mathbf{p}^0\alpha^0\rangle/(E' - E_k - a + ib). \tag{52}$$

Observemos que a y b son reales, y que b es positivo.

Este valor de $\langle k|H'\rangle$ sustituido en (49) nos da un coeficiente de dispersión

$$\frac{4\pi^2 h^2 W^0 W' P'}{c^4 P^0} \; \frac{|\langle \mathbf{p}'\alpha'|V|k\rangle|^2 |\langle k|V|\mathbf{p}^0\alpha^0\rangle|^2}{(E' - E_k - a)^2 + b^2}. \tag{53}$$

Para hallar la sección eficaz sobre la que debe incidir la partícula para ser dispersada de cualquier forma, integramos (53) para todas las direcciones de dispersión, es decir, para todas las direcciones del vector $\mathbf{p}'$ cuyo módulo P' es fijo, y luego sumamos para todos los α' que hemos de tener en consideración, o sea, aquellos para los que $W' > mc^2$. Con ayuda de (51), estos cálculos nos dan el resultado

$$\frac{4\pi h^2 W^0}{c^2 P^0} \; \frac{b|\langle k|V|\mathbf{p}^0\alpha^0\rangle|^2}{(E' - E_k - a)^2 + b^2}. \tag{54}$$

Si ahora variamos E' continuamente pasando por el valor E_k, la principal variación de (53) o (54) será debida al pequeño denominador $(E' - E_k - a)^2 + b^2$. Si despreciamos la dependencia de E' a través de los otros factores que figuran en (53) y (54), la máxima dispersión se obtendrá cuando E' tenga el valor $E_k + a$, y cuando E difiera de dicho valor en la cantidad b, la dispersión será la mitad de su valor máximo. La gran dispersión que aparece para valores de la energía de la partícula para los que E' es próximo a E_k da lugar al fenómeno de que aparezca una línea de absorción. El centro de dicha línea estará desplazado de la energía de resonancia de la partícula — o sea aquella energía que haría que la energía total fuera exactamente igual a E_k — en una cantidad a, en tanto que la cantidad b representa la llamada semiamplitud de la línea.

53. *Emisión y absorción*

Para el estudio de la emisión y absorción hemos de considerar estados no estacionarios y emplear el método de perturbaciones de § 44. Si queremos calcular el coeficiente de emisión espontánea hemos de tomar un estado inicial en el que la partícula esté absorbida, que corresponderá, por tanto, a un ket $|k\rangle$, y determinar la probabilidad de que en un instante posterior la partícula escape hacia infinito con un momento definido. Podemos aplicar el método de § 46. Según el resultado (39) de dicha sección, la probabilidad por unidad de tiempo y por unidad de intervalo de ω y de χ, de que la partícula sea emitida en cualquier dirección ω', χ' quedando el dispersor en el estado α' es

$$2\pi\hbar^{-1}|\langle W'\omega'\chi'\alpha'|V|k\rangle|^2, \tag{55}$$

siempre que α' sea tal que la energía W' de la partícula dada por (18) sea mayor que mc^2. Para valores de α' que no satisfagan dicha condición no hay emisión posible. Los elementos de matriz $\langle W'\omega'\chi'\alpha'|V|k\rangle$ deben referirse a una representación en la que W, ω, χ y α sean diagonales con la función de peso unidad. En cambio, los elementos de matriz de V que aparecen en las tres secciones anteriores se refieren a una representación en la que p_x, p_y, p_z son diagonales con la función de peso unidad, o en la que P, ω, χ son diagonales con la función de peso P^2 sen ω. Se refieren, por tanto, a una representación en la que W, ω, χ son diagonales con una función de peso $dP/dW \cdot P^2$ sen $\omega = WP/c^2 \cdot$ sen ω. Luego, el elemento de matriz $\langle W'\omega'\chi'\alpha'|V|k\rangle$ de (55) es igual a $(W'P'/c^2 \text{ sen } \omega')^{\frac{1}{2}}$ multiplicado por el elemento de matriz de las secciones anteriores $\langle W'\omega'\chi'\alpha'|V|k\rangle$ o $\langle \mathbf{p}'\alpha'|V|k\rangle$; de donde (55) es igual a

$$\frac{2\pi}{\hbar}\frac{W'P'}{c^2} \text{ sen } \omega'|\langle \mathbf{p}'\alpha'|V|k\rangle|^2.$$

Luego, la probabilidad de emisión por unidad de ángulo sólido y por unidad de tiempo de modo que el dispersor pase simultáneamente al estado α' será

$$\frac{2\pi}{\hbar}\frac{W'P'}{c^2}|\langle \mathbf{p}'\alpha'|V|k\rangle|^2. \tag{56}$$

Para hallar la probabilidad total por unidad de tiempo de que la partícula sea emitida en una dirección cualquiera dejando al dispersor en un estado final cualquiera, habremos de integrar (56) para todos los ángulos ω' y χ', y sumar después para todos los estados α' cuya energía $H_s(\alpha')$ verifique la condición $H_s(\alpha') + mc^2 < E_k$. El resultado da $2b/\hbar$, donde b es la cantidad definida en (51). *Por tanto, existe esta sencilla relación entre el coeficiente de emisión total y la semiamplitud de la línea de absorción.*

Consideremos ahora la absorción. Habremos de tomar un estado inicial para el cual la partícula no esté absorbida sino que incida con un momento definido. Luego, el ket que corresponde al estado inicial en este caso será de la forma (19). Debemos determinar la probabilidad de que la partícula sea absorbida después de un tiempo t. Dado que nuestro estado final k no pertenece a un intervalo de estados continuo, no podemos emplear directamente el resultado (39) de § 46. Pero si tomamos como ket correspondiente al estado inicial

$$|0\rangle = |\mathbf{p}^0\alpha^0\rangle, \tag{57}$$

sigue siendo aplicable el análisis de §§ 44 y 46 hasta la ecuación (36), y nos dice que la probabilidad de que la partícula sea absorbida para formar el estado k después de un tiempo t es

$$2|\langle k|V|\mathbf{p}^0\alpha^0\rangle|^2[1-\cos\{(E_k-E')t/\hbar\}]/(E_k-E')^2.$$

Este valor corresponde a una distribución de partículas de densidad h^{-3}, debido a la omisión del factor $h^{\frac{3}{2}}$ en (57) comparada con (19). Luego, la probabilidad de que la partícula sea absorbida tras un tiempo t cuando el número de partículas incidentes que atraviesan la unidad de área en la unidad de tiempo es la unidad vale

$$2h^3W^0/c^2P^0\cdot|\langle k|V|\mathbf{p}^0\alpha^0\rangle|^2[1-\cos\{(E_k-E')t/\hbar\}]/(E_k-E')^2. \tag{58}$$

Para obtener el coeficiente de absorción hemos de considerar partículas incidentes que no tengan todas exactamente el mismo valor de la energía $W^0 = E' - H_s(\alpha^0)$, sino con una distribución de los valores de la energía en torno al valor correcto $E_k - H_s(\alpha^0)$ requerido para que haya absorción. Si tomamos un haz de partículas incidentes que contenga una partícula por unidad de intervalo de energía y por unidad de tiempo que atraviesa la unidad de área, la probabilidad de que haya una absorción tras un tiempo t vendrá dada por la integral de 58 respecto a E'. Dicha integral puede ser calculada del mismo modo que (37) de § 46, y es igual a

$$4\pi^2h^2W^0t/c^2P^0\cdot|\langle k|V|\mathbf{p}^0\alpha^0\rangle|^2.$$

Luego, la probabilidad de absorción por unidad de tiempo para un haz incidente que contenga una partícula por unidad de intervalo de energía que atraviesa la unidad de área por unidad de tiempo será

$$4\pi^2h^2W^0/c^2P^0\cdot|\langle k|V|\mathbf{p}^0\alpha^0\rangle|^2, \tag{59}$$

que es el coeficiente de absorción.

Veamos cuál es la relación entre los coeficientes de absorción y de emisión (59) y (56), y los coeficientes de dispersión de resonancia calculados en la sección anterior. Cuando el haz incidente no está constituido por

partículas con la misma energía, sino por una distribución que contenga una partícula por unidad de intervalo de energía y de tiempo que atraviesa la unidad de área, el número total de partículas con energías próximas a una línea de absorción que serán dispersadas vendrá dado por la integral de (54) respecto a E'. Si despreciamos la dependencia del numerador de (54) respecto a E', dicha integral tendrá precisamente el valor (59), puesto que

$$\int_{-\infty}^{\infty} \frac{b}{(E' - E_k - a)^2 + b^2}\, dE' = \pi,$$

Luego, *el número total de partículas dispersadas en el entorno de una línea de absorción es igual al número total de partículas absorbidas.* Por lo tanto, podemos considerar que dichas partículas dispersadas han sido absorbidas y después reemitidas en otra dirección. Además, el número de partículas dispersadas por unidad de ángulo sólido en una dirección, dada por $\mathbf{p}'$, que pertenecen a dispersores en el estado α' en el entorno de una línea de absorción viene dado por la integral de (53) respecto a E', que calculada como antes tiene el valor

$$\frac{4\pi^2 h^2 W^0 W' P'}{c^4 P^0} \frac{\pi}{b} |\langle \mathbf{p}'\alpha' | V | k \rangle|^2 |\langle k | V | \mathbf{p}^0 \alpha^0 \rangle|^2.$$

Este valor es precisamente igual al coeficiente de absorción (59) multiplicado por el coeficiente de emisión (56) y dividido por el coeficiente de emisión total $2b/\hbar$. Esto concuerda con el punto de vista que considera las partículas dispersadas por resonancia como aquellas partículas que han sido absorbidas y después reemitidas, teniendo lugar ambos procesos de absorción y emisión cada uno con su propia ley probabilística independiente del otro, ya que según dicho punto de vista, la fracción del número total de partículas absorbidas que sería reemitida en una dirección dada por unidad de ángulo sólido, sería precisamente igual al coeficiente de emisión en dicha dirección dividido por el coeficiente de emisión total.

IX

SISTEMAS QUE CONTIENEN PARTÍCULAS IDÉNTICAS

54. *Estados simétricos y antisimétricos*

En física atómica si un sistema contiene varias partículas de la misma clase — por ejemplo, un conjunto de electrones —, dichas partículas son absolutamente indistinguibles entre sí. Al intercambiar dos de ellas no se produce ningún cambio observable. Esta circunstancia da lugar a fenómenos curiosos en mecánica cuántica, sin análogo en la teoría clásica, que derivan del hecho de que en mecánica cuántica pueden tener lugar transiciones en las que se intercambian dos partículas idénticas, y que no pueden ser detectadas por ningún procedimiento de observación. Evidentemente una teoría aceptable debería considerar que dos estados indistinguibles en toda observación son el mismo estado y que el intercambio de dos partículas idénticas no constituye ninguna transición. Más adelante veremos que es posible volver a formular la teoría de forma que verifique estas condiciones.

Sea un sistema que contiene n partículas idénticas. Como variables dinámicas podemos tomar un conjunto de variables ξ_1 para la primera partícula, el conjunto correspondiente ξ_2 para la segunda y así sucesivamente hasta la n-ésima. Con esta elección las ξ_r conmutarán con las ξ_s para $r \neq s$. (Puede ocurrir que sean necesarias otras variables extra para dar cuenta de otros componentes del sistema además de las n partículas, pero no será necesario mencionarlas explícitamente en este capítulo.) El hamiltoniano que describe el movimiento del sistema se podrá expresar como función de las $\xi_1, \xi_2, \ldots, \xi_n$. El hecho de que las partículas sean idénticas exige que el *hamiltoniano sea una función simétrica de* $\xi_1, \xi_2, \ldots, \xi_n$, es decir, que permanezca invariante cuando se intercambian o permutan de cualquier forma los conjuntos de variables ξ_r. Esta condición debe ser válida para cualquier perturbación aplicada al sistema. De hecho, cualquier cantidad que tenga significado físico debe ser una función simétrica de las ξ_r.

Sean $|a_1\rangle, |b_1\rangle, \ldots$, kets de la primera partícula considerada ella sola como un sistema dinámico. Existirán kets correspondientes $|a_2\rangle, |b_2\rangle, \ldots$, para la segunda partícula y así sucesivamente. Podemos construir un ket

ξ_1 ξ_2 ξ_2 ξ_1

para el conjunto mediante el producto de kets, cada uno de los cuales corresponde a una partícula considerada ella sola como un sistema; por ejemplo,

$$|a_1\rangle|b_2\rangle|c_3\rangle\ldots|g_n\rangle = |a_1b_2c_3\ldots g_n\rangle \tag{1}$$

según la notación (65) de § 20. El ket (1) corresponde a un tipo de estado particular del conjunto, que se caracteriza porque en él cada partícula está en su propio estado, simbolizado por el factor correspondiente que figura en el primer miembro de (1). El ket general del conjunto será una suma o integral de kets del tipo (1), y corresponde a un estado para el cual no se puede afirmar que cada partícula esté en su propio estado, sino únicamente que cada partícula está parcialmente en los distintos estados de forma relacionada con el modo cómo las otras partículas están parcialmente en los diversos estados. Si $|a_1\rangle$, $|b_1\rangle$, ..., es un conjunto de kets básicos para la primera partícula considerada sola, los kets $|a_2\rangle$, $|b_2\rangle$, ... constituirán a su vez un conjunto de kets básicos para la segunda partícula considerada ella sola y así sucesivamente; y en ese caso los kets (1) constituirán un conjunto de kets básicos para el conjunto. Diremos que la representación del conjunto formada por estos kets básicos es una *representación simétrica*, pues trata del mismo modo a todas las partículas.

En (1) podemos intercambiar los kets de las dos primeras partículas y obtener un nuevo ket del conjunto

$$|b_1\rangle|a_2\rangle|c_3\rangle\ldots|g_n\rangle = |b_1a_2c_3\ldots g_n\rangle.$$

De forma más general podemos intercambiar el papel de las dos primeras partículas de cada ket del conjunto y obtener nuevos kets del conjunto. El proceso de intercambiar las dos primeras partículas equivale a la aplicación de un operador a los kets del conjunto, que como los estudiados en § 7, será también lineal. Análogamente, el proceso de intercambiar cualquier par de partículas corresponde a un operador lineal, y por aplicación sucesiva de tales intercambios podemos obtener cualquier permutación de las partículas; por tanto, una permutación aparece ahora como la aplicación de un nuevo operador lineal a los kets del conjunto. Una *permutación* se dice que es *par* o *impar* según que sea necesario un número par o impar de intercambios para formarla.

Diremos que un ket $|X\rangle$ del conjunto es *simétrico* si es invariante al aplicarle cualquier permutación, es decir si

$$P|X\rangle = |X\rangle \tag{2}$$

para cualquier permutación P. Y diremos que es *antisimétrico* si no cambia al aplicarle una permutación par y cambia de signo al aplicarle una permutación impar, es decir, si

$$P|X\rangle = \pm|X\rangle, \tag{3}$$

* o estados con grados de libertad periódicos

donde los signos + o — dependen de si P es par o impar. Los estados correspondientes a kets simétricos se denominan *estados simétricos* y los que corresponden a kets antisimétricos, *estados antisimétricos*.* En una representación simétrica el representante de un ket simétrico será una función simétrica de las variables de las distintas partículas, y el representante de un ket antisimétrico será una función antisimétrica.

En la imagen de Schrödinger el ket que corresponde a un estado del conjunto variará con el tiempo según la ecuación de movimiento de Schrödinger. Si inicialmente era un ket simétrico deberá permanecer siempre simétrico, pues, teniendo en cuenta que el hamiltoniano es simétrico, no hay ningún elemento que pueda destruir o cambiar la simetría. Análogamente si el ket era inicialmente antisimétrico deberá permanecer siempre antisimétrico. Luego, *todo estado que inicialmente es simétrico seguirá siéndolo siempre, y todo estado que inicialmente es antisimétrico conservará igualmente su antisimetría.* En consecuencia, puede ocurrir que para una determinada clase de partículas no se den en la naturaleza más que estados simétricos o antisimétricos. Si ocurriese alguna de estas dos posibilidades, daría lugar a ciertos fenómenos especiales para las partículas en cuestión.

Supongamos en primer lugar que únicamente se presenten en la naturaleza estados antisimétricos. El ket (1) no es antisimétrico y, por tanto, no corresponde a ningún estado real. A partir de (1) podemos construir siempre un ket antisimétrico aplicándole todas las posibles permutaciones y sumando los resultados, después de multiplicar los términos que provienen de una permutación impar por —1. Así obtenemos

$$\sum_P \pm P|a_1 b_2 c_3 \ldots g_n\rangle, \tag{4}$$

en donde el signo + o — depende de si P es par o impar. El ket (4) se puede escribir en forma de determinante

$$\begin{vmatrix} |a_1\rangle & |a_2\rangle & |a_3\rangle & \cdot & \cdot & \cdot & |a_n\rangle \\ |b_1\rangle & |b_2\rangle & |b_3\rangle & \cdot & \cdot & \cdot & |b_n\rangle \\ |c_1\rangle & |c_2\rangle & |c_3\rangle & \cdot & \cdot & \cdot & |c_n\rangle \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ |g_1\rangle & |g_2\rangle & |g_3\rangle & \cdot & \cdot & \cdot & |g_n\rangle \end{vmatrix} \tag{5}$$

y su representante en una representación simétrica será un determinante. El ket (4) o (5) no es el ket antisimétrico general sino únicamente un simple ket particular. Corresponde a un estado del conjunto en el que están ocupados ciertos estados-partícula, a saber, los estados a, b, c, ..., g, no pudiendo decirse qué partícula está en cada estado, pues cualquier partícula tiene la misma probabilidad de ocupar cualquiera de dichos estados. Si dos

de los estados-partícula a, b, c, ..., g son el mismo, el ket (4) o (5) es nulo y no corresponde a ningún estado del conjunto. Luego, *dos partículas no pueden ocupar el mismo estado.* Más en general, *los estados ocupados han de ser todos independientes,* pues sino (4) o (5) son nulos. Ésta es una característica importante de las partículas que sólo pueden ocupar estados antisimétricos. Ello da lugar a una estadística especial, que fue estudiada por vez primera por Fermi, y así llamaremos *fermiones* a las partículas que sólo pueden ocupar estados antisimétricos.

Supongamos ahora que únicamente se presenten estados simétricos. El ket (1) no es simétrico a menos que todos los estados-partícula a, b, c, ..., g sean el mismo estado, pero siempre será posible construir a partir de él un estado simétrico aplicándole todas las permutaciones posibles y sumando los resultados; así se obtiene

$$\sum_P P|a_1 b_2 c_3 \ldots g_n\rangle. \tag{6}$$

El ket (6) no es el ket simétrico general, sino un ket particular sencillo. Corresponde a un estado del conjunto para el que están ocupados ciertos estados-partícula, a saber, los estados a, b, c, ..., g, no pudiendo decirse qué partícula está en cada estado. En este caso es posible que dos o más estados a, b, c, ..., g sean el mismo estado. Como consecuencia de ello, la estadística a que obedecerán las partículas no será la misma que la que se utiliza corrientemente en la teoría clásica. Esta nueva estadística fue estudiada por primera vez por Bose, y así llamaremos *bosones* a las partículas que sólo pueden ocupar estados simétricos.

Podemos poner de manifiesto la diferencia entre la estadística de Bose y la estadística ordinaria considerando el caso particular en que no haya más que dos partículas y dos estados independientes a y b para una de ellas. Según la mecánica clásica, si el conjunto de las dos partículas está en equilibrio termodinámico a temperatura elevada, cada partícula tendrá la misma probabilidad de estar en cualquiera de los dos estados. Por tanto, habrá una probabilidad $\frac{1}{4}$ de que ambas partículas estén en el estado a, una probabilidad $\frac{1}{4}$ de que ambas estén en el estado b, y una probabilidad $\frac{1}{2}$ de que una partícula esté en un estado y la otra en el otro. En la teoría cuántica existen tres estados simétricos independientes para el par de partículas, que corresponden a los kets simétricos $|a_1\rangle|a_2\rangle$, $|b_1\rangle|b_2\rangle$, y $|a_1\rangle|b_2\rangle + |a_2\rangle|b_1\rangle$, y que representan respectivamente los estados en que ambas partículas están en el estado a, ambas en el estado b y una partícula en un estado y la otra en el otro. En equilibrio termodinámico a temperatura elevada, los tres estados serán igualmente probables como se vio en § 33, y, en consecuencia, existirá una probabilidad $\frac{1}{3}$ de que ambas partículas estén en el estado a, una probabilidad $\frac{1}{3}$ de que ambas estén en el estado b, y una probabilidad $\frac{1}{3}$ de que una partícula esté en un estado y la otra en el otro. Luego, *en la estadística de Bose la probabilidad de que*

dos partículas estén en el mismo estado es mayor que en la estadística clásica. La estadística de Bose difiere de la estadística clásica en sentido contrario al que difiere la de Fermi, para la cual la probabilidad de que dos partículas ocupen el mismo estado es nula.

Al establecer una teoría de los átomos siguiendo el camino indicado al comienzo de § 38, es necesario, a fin de explicar los resultados experimentales, suponer que dos electrones nunca pueden estar en el mismo estado. Esta regla se conoce con el nombre de *principio de exclusión de Pauli.* Esto nos demuestra que los *electrones son fermiones.* La ley de Planck de la radiación nos dice que los fotones son bosones, pues únicamente la estadística de Bose conduce a dicha ley. Análogamente, existen razones experimentales para todas las demás partículas conocidas en física que indican que son fermiones o bosones. Los protones, neutrones y positrones son fermiones, mientras que las partículas α son bosones. Todas las partículas que se dan en la naturaleza son o bien fermiones o bien bosones, y, en consecuencia, los únicos estados que se presentan en la práctica para un conjunto de partículas idénticas son o simétricos o antisimétricos. Matemáticamente son posibles otras clases de simetría más complicada; pero no se presentan para ninguna de las partículas conocidas. En una teoría en la que una misma partícula admite únicamente estados simétricos o antisimétricos, no hay ninguna distinción entre dos estados que sólo difieren en la ordenación de las partículas, y así las transiciones mencionadas al comenzar esta sección dejan de existir.

55. *Las permutaciones como variables dinámicas*

Vamos a construir una teoría general aplicable a los sistemas que contienen n partículas idénticas cuyos estados pueden presentar cualquier clase de simetría, es decir, sin imponer ninguna restricción de que sólo puedan darse estados simétricos o sólo estados antisimétricos. El estado general no será simétrico ni antisimétrico, y para $n > 2$, ni siquiera se podrá expresar linealmente en función de estados simétricos y antisimétricos. Esta teoría no se puede aplicar directamente a ninguna clase de partículas existente, no obstante, será útil para obtener un método aproximado de tratar un conjunto de electrones como veremos en § 58.

Hemos visto que cualquier permutación P de las n partículas es un operador lineal que puede ser aplicado a cualquier ket del conjunto. Luego, P puede ser considerada como una variable dinámica de nuestro sistema de n partículas. Existen $n!$ permutaciones, y cada una de ellas puede considerarse como una variable dinámica. Entre ellas está la permutación idéntica P_1 que es igual a la unidad. El producto de dos permutaciones cualesquiera es otra permutación, y por tanto, cualquier función

de permutaciones puede reducirse a una función lineal. Toda permutación P tiene su inversa P^{-1} que verifica

$$PP^{-1} = P^{-1}P = P_1 = 1.$$

La permutación P se puede aplicar al bra $\langle X|$ del conjunto, obteniéndose un nuevo bra, que de momento designaremos por $P\langle X|$. Si aplicamos P a ambos factores del producto $\langle X|Y\rangle$, éste debe quedar inalterado por ser dicho producto un número que no depende del orden de las partículas. Luego,

$$(P\langle X|)P|Y\rangle = \langle X|Y\rangle$$

de donde

$$P\langle X| = \langle X|P^{-1} \tag{7}$$

Pero $P\langle X|$ es el conjugado imaginario de $P|X\rangle$ y, por tanto, es igual a $\langle X|\bar{P}$, luego, según (7) se tiene

$$\bar{P} = P^{-1}. \tag{8}$$

En consecuencia, las permutaciones no son en general variables dinámicas reales, siendo el complejo conjugado de una permutación igual a su recíproco.

Cualquier permutación de los números 1, 2 ,3, ..., n se puede expresar como producto de ciclos; por ejemplo para $n = 8$ se tiene

$$P_a = (143)(27)(58)(6), \tag{9}$$

donde dentro de cada paréntesis todo número debe ser reemplazado por el siguiente menos el último que debe ser sustituido por el primero. Así pues, P_a cambia los números 12345678 por 47138625. La clase de permutación queda determinada por la partición del número n dada por la cantidad de números que hay en cada paréntesis. Así la clase de P_a viene dada por la partición $8 = 3 + 2 + 2 + 1$. Diremos que las permutaciones de la misma clase, es decir, correspondientes a la misma partición son permutaciones *equivalentes*. Así por ejemplo, P_a de (9) es equivalente a

$$P_b = (871)(35)(46)(2). \tag{10}$$

El conjunto de las $n!$ permutaciones posibles puede dividirse en conjuntos de permutaciones equivalentes, y cada uno de dichos conjuntos se denomina una *clase* de equivalencia. La permutación $P_1 = 1$ constituye por sí sola una clase. Toda permutación es equivalente a su recíproca.

Cuando dos permutaciones P_a y P_b son equivalentes, cualquiera de ellas, por ejemplo P_b, se puede obtener mediante una permutación P_x de la otra P_a. Así, para las permutaciones (9) y (10) podemos tomar como P_x la permutación que cambia 14327586 por 87135462, es decir, la permutación

$$P_x = (18623)(475).$$

Los distintos modos de escribir P_a y P_b en notación cíclica darían lugar a distintas P_x. Cualquiera de dichas P_x aplicada al producto $P_a|X\rangle$ daría como resultado $P_b \cdot P_x|X\rangle$, es decir,

$$P_x P_a|X\rangle = P_b P_x|X\rangle.$$

Luego,

$$P_b = P_x P_a P_x^{-1}, \tag{11}$$

que expresa la condición para que P_a y P_b sean equivalentes en forma de ecuación algebraica. La existencia de un P_x cualquiera que verifique (11) es suficiente para asegurar que P_a y P_b son equivalentes.

56. *Las permutaciones como constantes del movimiento*

Toda función simétrica V de las variables dinámicas de todas las partículas queda inalterada cuando se le aplica una permutación cualquiera P, y por tanto, al aplicar P al producto $V|X\rangle$ queda aplicada al factor $|X\rangle$, o sea,

$$PV|X\rangle = VP|X\rangle.$$

Luego,

$$PV = VP, \tag{12}$$

que nos viene a decir que *toda función simétrica de las variables dinámicas conmuta con las permutaciones.* El hamiltoniano es una función simétrica de las variables dinámicas, y por tanto conmuta con cualquier permutación. Luego, *toda permutación es una constante del movimiento.* Este resultado es válido incluso cuando el hamiltoniano no es constante. Si $|Xt\rangle$ es una solución cualquiera de la ecuación de movimiento de Schrödinger, $P|Xt\rangle$ es también solución.

En mecánica cuántica, cuando estudiamos un sistema y encontramos una constante del movimiento α, sabemos que si en un estado de movimiento cualquiera tenía inicialmente el valor numérico α', seguirá teniendo siempre dicho valor, pudiendo así asignar distintos valores numéricos α' a los distintos estados y de este modo clasificarlos. Pero cuando tenemos distintas constantes del movimiento α que no conmutan entre sí (como ocurre con las permutaciones P), el procedimiento no es tan directo, pues, en general, no podemos asignar valores numéricos simultáneamente a todas las α en ningún estado. Consideremos en primer lugar el caso de un sistema cuyo hamiltoniano no es función explícita del tiempo. La existencia de constantes del movimiento α que no conmutan entre sí indica que el sistema es degenerado. En efecto, para un sistema no degenerado, el hamiltoniano H constituye por sí solo un conjunto completo de observables que

conmutan, y en consecuencia, según el teorema 2 de § 19, cualquier α es función de H, y así H conmuta con cualquier α.

Hemos de buscar una función β de las α que para todos los estados que pertenezcan a un mismo nivel de energía H', tenga el mismo valor numérico β', pudiendo de esta forma utilizar β para clasificar los niveles de energía del sistema. La condición que ha de verificar β es la de ser una función de H y, por tanto, ha de conmutar con toda variable dinámica que conmute con H, es decir, con toda constante del movimiento. Si las α son las únicas constantes del movimiento, o bien si existe un conjunto de α que conmuta con todas las demás constantes del movimiento independientes, nuestro problema se reduce a hallar una función β de las α que conmute con todas las α. Con ello podremos asignar a β un valor numérico β' para cada nivel de energía del sistema. Si existen distintas funciones β, han de conmutar unas con otras a fin de que podamos asignar simultáneamente a todas ellas valores numéricos. Obtenemos así una clasificación de los niveles de energía. Cuando el hamiltoniano es función explícita del tiempo, no podemos hablar de niveles de energía, pero las β seguirán siendo una clasificación útil de los niveles de energía.

Vamos a seguir este procedimiento para estudiar las permutaciones P. Hemos de hallar una función χ de las P tal que $P\chi P^{-1} = \chi$ para toda P. Evidentemente una χ posible será $\sum P_c$, la suma de todas las permutaciones de una clase c, es decir, la suma del conjunto de todas las permutaciones equivalentes, ya que $\sum PP_cP^{-1}$ estará formado por las mismas permutaciones sumadas en un orden distinto. Existirá una de estas χ para cada clase. Además, no puede existir ninguna otra χ independiente, pues toda función de las P puede expresarse como función lineal de ellas con coeficientes numéricos, y no conmutarán con todas las P a menos que los coeficientes de las P equivalentes sean idénticos. Así obtenemos todas las χ que sirven para clasificar los estados. Es conveniente definir cada χ como una media y no como una suma, y así

$$\chi_c = n_c^{-1} \sum P_c,$$

siendo n_c el número de P que hay en cada clase c. Otra forma de escribir χ_c es

$$\chi_c = n!^{-1} \sum_P PP_cP^{-1}, \tag{13}$$

donde la suma está extendida a todas las $n!$ permutaciones P, siendo fácil ver que dicha suma contiene cada elemento de la clase c el mismo número de veces. Para cada permutación P existe una χ, que podemos llamar $\chi(P)$, igual a la media de todas las permutaciones equivalentes a P. Una de las χ es $\chi(P_1) = 1$.

Las constantes del movimiento $\chi_1, \chi_2, \ldots, \chi_m$ así obtenidas tendrán un valor numérico definido para cada estado estacionario del sistema cuando

el hamiltoniano no sea función explícita del tiempo, y en el caso general pueden servir asimismo para clasificar los estados, pues existirá un conjunto de estados para cada conjunto de valores permitidos $\chi'_1, \chi'_2, \ldots, \chi'_m$ de las χ. Dado que las χ son constantes del movimiento en todos los casos, dichos conjuntos de estados serán *exclusivos*, es decir, que nunca tendrán lugar transiciones desde un estado de un conjunto a un estado de otro conjunto.

Los conjuntos de valores posibles de χ' que se pueden asignar a las χ están limitados por el hecho de que existen relaciones algebraicas entre las χ. El producto de dos χ cualesquiera $\chi_p \chi_q$ se puede expresar linealmente en función de las P, y puesto que conmuta con todas las P tiene que poderse expresar como función lineal de las χ; tendremos pues

$$\chi_p \chi_q = a_1 \chi_1 + a_2 \chi_2 + \ldots + a_m \chi_m, \tag{14}$$

donde las a_i son números. Cualesquiera valores numéricos χ' que se asignen a las χ han de ser autovalores suyos y han de satisfacer estas mismas ecuaciones algebraicas. Para cada solución χ' de estas ecuaciones existe un conjunto exclusivo de estados. Una solución evidente es $\chi'_p = 1$ para toda χ_p, que corresponde al conjunto de estados simétricos. Otra solución también obvia, que da lugar al conjunto de estados antisimétricos es $\chi'_p = \pm 1$, con signo más o signo menos según sean pares o impares las permutaciones de la clase p. Las demás soluciones pueden ser halladas en cualquier caso particular con los métodos algebraicos ordinarios, pues los coeficientes a_i de (14) pueden obtenerse directamente considerando los tipos de permutación a que hacen referencia las χ que intervienen. Cada solución es, salvo un factor numérico, lo que en teoría de grupos se denomina un *carácter* del grupo de las permutaciones. Las χ son todas variables dinámicas reales, puesto que cada P y su compleja conjugada P^{-1} son equivalentes y en la definición de las χ entrarán, en consecuencia, sumadas; por tanto las χ' han de ser números reales.

El número de soluciones posibles de las ecuaciones (14) puede ser calculado fácilmente, pues tiene que ser igual al número de autovalores distintos de una función arbitraria B de las χ. Con ayuda de las ecuaciones (14) podremos expresar B linealmente en función de las χ; y así

$$B = b_1 \chi_1 + b_2 \chi_2 + \ldots + b_m \chi_m. \tag{15}$$

Análogamente podemos expresar cada una de las cantidades $B^2, B^3, \ldots, B^m$ como función lineal de las χ. De las m ecuaciones así obtenidas y la ecuación $\chi(P_1) = 1$ podemos eliminar las m incógnitas $\chi_1, \chi_2, \ldots, \chi_m$ obteniendo como resultado una ecuación algebraica de grado m en B,

$$B^m + c_1 B^{m-1} + c_2 B^{m-2} + \ldots + c_m = 0.$$

Las m soluciones de esta ecuación determinan los m autovalores posibles de B, cada uno de los cuales será una función lineal de las $b_1, b_2, \ldots, b_m$ según (15), cuyos coeficientes serán un conjunto de posibles valores $\chi'_1, \chi'_2, \ldots, \chi'_m$. Los conjuntos de valores χ' así obtenidos han de ser todos distintos, pues si hubiera menos de m conjuntos de valores χ' de las χ distintos, existiría una función lineal de las χ cuyos autovalores serían todos nulos, lo que implicaría que la propia función fuese idénticamente nula, y las χ no serían linealmente independientes. Luego, el número de conjuntos de valores numéricos de las χ es exactamente igual a m, o sea, igual al número de clases de permutaciones que a su vez es igual al número de particiones de n. En consecuencia, el número de conjuntos exclusivos de estados es igual a dicho número.

Todas las variables dinámicas que tienen importancia física y todas las cantidades observables son simétricas respecto a las partículas, y por tanto, conmutan con todas las P. Luego, las únicas funciones de las P que tienen importancia física son las χ. Los estados que corresponden a $|\chi'\rangle$ y a $f(P)|\chi'\rangle$, donde $|\chi'\rangle$ es un autoket de las χ perteneciente a los autovalores χ' y $f(P)$ es una función cualquiera de las P tal que $f(P)|\chi'\rangle \neq 0$, son indistinguibles para toda observación y, por tanto, son equivalentes desde el punto de vista físico. Multiplicando $|\chi'\rangle$ por funciones de las P podremos formar un número determinado $n(\chi')$ de kets independientes, que dependerá únicamente de las χ'. Dicho número será igual al número de filas o columnas de una representación matricial de las P en la que cada χ sea igual a χ'. Si $|\chi'\rangle$ corresponde a un estado estacionario, $n(\chi')$ será un grado de degeneración (en lo que concierne a la degeneración producida por la simetría entre las partículas). Esta degeneración no puede ser eliminada por ninguna perturbación simétrica respecto a las partículas.

57. *Determinación de los niveles de energía*

Apliquemos el método de perturbaciones de § 43 para calcular en primera aproximación los niveles de energía, en el caso de que el hamiltoniano no sea función explícita del tiempo. Supondremos que para los estados estacionarios del conjunto del sistema no perturbado, cada una de las partículas idénticas está en su propio estado. Para n partículas tendremos n estados que corresponderán a los kets $|\alpha^1\rangle$, $|\alpha^2\rangle$, ..., $|\alpha^n\rangle$, que de momento supondremos ortogonales. El ket del conjunto será en este caso

$$|X\rangle = |\alpha_1^1\rangle|\alpha_2^2\rangle\ldots|\alpha_n^n\rangle, \tag{16}$$

igual que (1) pero con $\alpha^1, \alpha^2, \ldots$, en lugar de $a, b, \ldots$ Aplicándole una permutación P obtenemos otro ket

$$P|X\rangle = |\alpha_r^1\rangle|\alpha_s^2\rangle\ldots|\alpha_z^n\rangle \tag{17}$$

donde $r, s, \ldots, z$ es una permutación de los números 1, 2, ..., n, que corresponderá a otro estado estacionario del conjunto con la misma energía. En conjunto habrá pues, $n!$ estados no perturbados con dicha energía, si suponemos que no hay ninguna otra causa de degeneración. Según el método de § 43, cuando el sistema no perturbado es degenerado hemos de considerar los elementos de matriz de la energía de perturbación V que hacen referencia a dos estados con la misma energía, es decir, los estados del tipo $\langle X|P_a V P_b|X\rangle$. Dichos elementos formarán una matriz de $n!$ filas y columnas, cuyos autovalores serán las correcciones de primer orden de los niveles de energía.

Hemos de definir otro tipo de operadores de permutación aplicable a los kets de la forma (17), que actúan sobre los índices de las α. Los operadores de permutación de este tipo los designaremos por P^α. La diferencia esencial entre los operadores P y los P^α puede ponerse de manifiesto del modo siguiente. Consideremos en general la permutación que cambia 2 por 3. Puede interpretarse bien como el intercambio de los objetos 2 y 3, o bien como el intercambio de los objetos que ocupan los lugares 2 y 3, siendo distinto el resultado en ambos casos. La primera de estas interpretaciones es la que corresponde a los operadores P, siendo los objetos en cuestión las partículas idénticas. Las permutaciones P pueden aplicarse a cualquier ket del conjunto. En cambio, la segunda interpretación de las permutaciones sólo tiene sentido cuando se aplica a kets de la forma (17), para los cuales cada partícula está en un lugar dado por α, o bien, cuando se aplica a una suma de dichos kets. Las permutaciones P pueden considerarse como variables dinámicas ordinarias, mientras que las permutaciones P^α sólo pueden considerarse como variables dinámicas en un sentido restringido, aplicable únicamente cuando se consideran estados que se pueden obtener por superposición de los distintos estados (17), que es lo que ocurre en nuestro problema de perturbación.

Podemos formar funciones algebraicas de las P^α que serán nuevos operadores aplicables a kets de la forma (17). En particular podemos construir $\chi(P_c^\alpha)$, o sea, la media de todos los P^α de una cierta clase. Dicha función ha de ser igual a $\chi(P_c)$, —promedio de los operadores de permutación P de la misma clase— pues el conjunto de todas las permutaciones de una clase tiene que ser a todas luces el mismo, tanto si las permutaciones se aplican a las partículas como si se aplican a los lugares que ocupan dichas partículas. Toda P conmuta con toda P^α, es decir,

$$P_a P_b^\alpha = P_b^\alpha P_a \tag{18}$$

Designando las α con los mismos números 1, 2, 3, ..., n que empleamos para designar las partículas, podemos establecer una correspondencia biyectiva entre las α y las partículas, y dada una permutación P_a cualquiera

aplicada a las partículas, dar significado a la misma permutación P_a^α aplicada a las α. Dicho significado es tal que para el ket $|X\rangle$ dado por (16)

$$P_a^\alpha P_a|X\rangle = |X\rangle. \tag{19}$$

Puesto que los kets $|\alpha^1\rangle$, $|\alpha^2\rangle$, ..., son ortogonales, $|X\rangle$ y $P|X\rangle$ serán también ortogonales, salvo para $P = 1$. Luego, para cualesquiera coeficientes c_P se verifica

$$\sum_P c_P\langle X|P^\alpha P_a|X\rangle = c_{P_a} \tag{20}$$

donde $|X\rangle$ ha de estar normalizado, y la sumación se extiende a todas las $n!$ permutaciones P o P^α dejando P_a fija. Definamos ahora V_P por

$$V_P = \langle X|VP|X\rangle. \tag{21}$$

Con ello resulta, para dos permutaciones P_x y P_y cualesquiera,

$$\begin{aligned}\langle X|P_xVP_y|X\rangle &= \langle X|VP_xP_y|X\rangle = V_{P_xP_y}\\ &= \sum_P V_P\langle X|P^\alpha P_xP_y|X\rangle\end{aligned}$$

con ayuda de (20). Mediante (18) se obtiene

$$\langle X|P_xVP_y|X\rangle = \sum_P V_P\langle X|P_xP^\alpha P_y|X\rangle. \tag{22}$$

Este resultado se puede escribir en la forma

$$V \approx \sum_P V_P\, P^\alpha, \tag{23}$$

donde el signo $\approx$ significa que la ecuación es válida en sentido restringido, es decir, que los operadores que figuran en ambos miembros son iguales mientras se utilicen con kets de la forma $P|X\rangle$ y sus bras imaginarios conjugados.

La fórmula (23) nos dice que la energía de perturbación V es igual, en sentido restringido, a una función lineal de los operadores de permutación P^α de coeficientes V_P dados por (21). Este sentido restringido es adecuado para calcular las correcciones de primer orden de los niveles de energía, pues dicho cálculo sólo lleva consigo elementos de matriz de V dados por (22). La fórmula (23) es muy útil dada la facilidad con que se opera con su segundo miembro.

Como ejemplo de aplicación de (23) vamos a determinar el valor medio de la energía para todos los estados que, perteneciendo a un conjunto exclusivo, provienen del estado no perturbado (16). Para ello hemos de calcular el valor medio de V para los estados (17) en que las χ tienen valores numéricos determinados χ'. Pero el valor medio de P_a^α en cualquiera de dichos estados es igual al de $P^\alpha P_a^\alpha(P^\alpha)^{-1}$, donde P^α es arbitrario, y por

tanto vale $n!^{-1} \sum_{P^\alpha} P^\alpha P_a^\alpha (P^\alpha)^{-1}$, o sea $\chi'(P_a^\alpha)$ o $\chi'(P_a)$. Luego, el valor medio de V es $\sum_P V_P \chi'(P)$. Podemos emplear un método análogo para calcular el valor medio de cualquier función de V, para lo cual únicamente hemos de sustituir cada P^α por $\chi'(P)$ en el cálculo del promedio.

El número de niveles de energía que hay en un conjunto exclusivo $\chi = \chi'$ que provienen de un estado dado del sistema no perturbado es igual al número de autovalores del segundo miembro de (23) que son compatibles con las ecuaciones $\chi = \chi'$. Dicho número no es más que el número $n(\chi')$ introducido al final de la sección precedente, y es igual al grado de degeneración de los estados de dicho conjunto.

Hemos supuesto que los kets individuales $|\alpha^1\rangle$, $|\alpha^2\rangle$, ..., que determinan los estados no perturbados mediante (16) son ortogonales. La teoría puede generalizarse fácilmente para el caso en que algunos de ellos estén repetidos, y los que son distintos sigan siendo ortogonales. En este caso tenemos algunas permutaciones P^α para las que $P^\alpha|X\rangle = |X\rangle$, que serán todas aquellas que consistan únicamente en intercambios de las α iguales. La ecuación (20) seguirá siendo válida si extendemos la suma únicamente a aquellas P que dan distintos valores de $P^\alpha|X\rangle$. Con este cambio en el significado de $\sum_P$, siguen siendo válidas todas las ecuaciones anteriores, incluso el resultado (23). Para el $|X\rangle$ de ahora existirán limitaciones sobre los posibles valores numéricos de las χ y así, por ejemplo, no pueden tomar los valores que corresponden a un $|X\rangle$ antisimétrico.

58. *Aplicación a electrones*

Consideremos el caso de que las partículas idénticas sean electrones. Según el principio de exclusión de Pauli discutido en § 54, únicamente habremos de considerar estados antisimétricos. Es necesario considerar explícitamente el hecho de que los electrones tienen spin, que se manifiesta por la aparición de un momento angular y un momento magnético. El efecto del spin en el movimiento de un electrón bajo la acción de un campo electromagnético no es muy grande. Sobre el electrón actúan unas fuerzas adicionales debidas a su momento magnético, que introducen nuevos términos en el hamiltoniano. El momento angular de spin no tiene ninguna acción directa sobre el movimiento, pero entra en juego cuando existen fuerzas que tienden a hacer girar su momento magnético, pues el momento magnético y el momento angular han de tener siempre la misma dirección. En ausencia de un campo magnético fuerte dichos efectos son pequeños, del mismo orden que las correcciones introducidas por la mecánica relativista, y no sería útil considerar dichos efectos en una teoría no relativista. La importancia del spin no reside en estos pequeños

efectos sobre el movimiento del electrón sino en el hecho de que proporciona dos estados internos del electrón, en correspondencia con los dos valores posibles de la componente del spin en una dirección cualquiera, y que hace que el número de estados independientes de un electrón sea doble. Este hecho, combinado con el principio de exclusión de Pauli, tiene consecuencias muy importantes.

Para un conjunto de electrones tenemos dos clases de variables dinámicas. Las primeras son las coordenadas x, y, z de todos los electrones y sus momentos canónicos conjugados p_x, p_y, p_z que denominaremos *variables orbitales.* Las segundas son las variables de spin σ_x, σ_y, σ_z de todos los electrones, introducidas en § 37. Ambas clases de variables pertenecen a grados de libertad distintos. Según §§ 20 y 21, existen kets que determinan el estado del sistema total de la forma $|A\rangle|B\rangle$, siendo $|A\rangle$ un ket que hace referencia únicamente a las variables orbitales y $|B\rangle$ un ket que únicamente hace referencia a variables de spin, y el ket general que determina el estado del sistema total será una suma o integral de kets de este tipo. Este modo de considerar el problema nos permite introducir dos clases de operadores de permutación, los P^x que actúan sobre las variables orbitales y que únicamente se aplican al factor $|A\rangle$ y los P^σ que actúan sobre las variables de spin y se aplican únicamente al factor $|B\rangle$. Los P^x y los P^σ podrán aplicarse a cualquier ket del sistema total, y no únicamente a ciertos kets particulares como ocurría con los P^α de la sección anterior. Las permutaciones P se aplicaban a todas las variables dinámicas de las partículas en cuestión, y por tanto para los electrones actuarán tanto sobre las variables orbitales como sobre las variables de spin. Así pues, cada P_a será igual al producto

$$P_a = P_a^x P_a^\sigma \tag{24}$$

Veamos ahora la necesidad de tener en cuenta las variables de spin al aplicar el principio de exclusión de Pauli, aunque despreciemos las fuerzas de spin en el hamiltoniano. P_a ha de tener el valor ± 1 en todo estado existente, según sea una permutación par o impar; luego, según (24)

$$P_a^x P_a^\sigma = \pm 1 \tag{25}$$

La teoría de las tres secciones precedentes resultaría trivial si se aplicase directamente a los electrones, para los cuales cada $P_a = \pm 1$. En cambio, podemos aplicarla a las permutaciones P^x de los electrones. Si despreciamos los términos del hamiltoniano que provienen de las fuerzas de spin, las P^σ son constantes del movimiento, ya que, en tal caso, las variables dinámicas de spin $\boldsymbol{\sigma}$ no figuran para nada en el hamiltoniano. Luego, las P^x también han de ser constantes del movimiento. Podemos, pues, introducir unas nuevas χ iguales a la media de todas las P^x de cada clase, y afirmar

que para cada conjunto permitido de valores numéricos χ' de dichas χ, existe un conjunto exclusivo de estados. Luego, existen conjuntos exclusivos de estados para los sistemas que contienen muchos electrones, pese a que nos limitemos a considerar únicamente estados que satisfacen el principio de exclusión de Pauli. Evidentemente, los conjuntos exclusivos de estados, lo serán ahora, únicamente de forma aproximada, pues las χ solamente son constantes del movimiento cuando despreciamos las fuerzas de spin. En este caso existirá una pequeña probabilidad de que tenga lugar una transición desde un estado de un conjunto a un estado de otro conjunto.

La ecuación (25) relaciona de un modo sencillo las P^x y las P^σ, y nos permite obtener todos los resultados que deseemos, como, por ejemplo, los caracteres χ', estudiando las variables dinámicas P^σ en lugar de las P^x. Las P^σ son mucho más fáciles de estudiar debido a que no existen más que dos estados de spin independientes para cada electrón. Este hecho se traduce en que haya menos caracteres χ' para el grupo de permutaciones de las variables σ que para el grupo general de las permutaciones, pues impide que un ket de las variables de spin sea antisimétrico respecto a más de dos de ellas.

El estudio de las P^σ es particularmente sencillo debido a que se pueden expresar como funciones algebraicas de las variables dinámicas $\boldsymbol{\sigma}$. Consideremos la cantidad

$$O_{12} = \tfrac{1}{2}\{1 + \sigma_{x1}\,\sigma_{x2} + \sigma_{y1}\,\sigma_{y2} + \sigma_{z1}\,\sigma_{z2}\} = \tfrac{1}{2}\{1 + (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)\}.$$

Con ayuda de las ecuaciones (50) y (51) de § 37 obtenemos fácilmente

$$(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)^2 = (\sigma_{x1}\,\sigma_{x2} + \sigma_{y1}\,\sigma_{y2} + \sigma_{z1}\,\sigma_{z2})^2 = 3 - 2(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2), \tag{26}$$

y en consecuencia,

$$O_{12}^2 = \tfrac{1}{4}\{1 + 2(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) + (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)^2\} = 1. \tag{27}$$

Además resulta

$$O_{12}\,\sigma_{x1} = \tfrac{1}{2}\{\sigma_{x1} + \sigma_{x2} - i\sigma_{z1}\,\sigma_{y2} + i\sigma_{y1}\sigma_{z2}\},$$
$$\sigma_{x2}\,O_{12} = \tfrac{1}{2}\{\sigma_{x2} + \sigma_{x1} + i\sigma_{y1}\,\sigma_{z2} - i\sigma_{z1}\sigma_{y2}\}$$

luego,

$$O_{12}\,\sigma_{x1} = \sigma_{x2}\,O_{12}.$$

Para σ_{y1} y σ_{z1} son válidas relaciones análogas

$$O_{12}\,\boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_2\,O_{12}$$

o sea,

$$O_{12}\,\boldsymbol{\sigma}_1\,O_{12}^{-1} = \boldsymbol{\sigma}_2.$$

De aquí, con ayuda de (27) podemos obtener

$$O_{12}\,\boldsymbol{\sigma}_2\,O_{12}^{-1} = \boldsymbol{\sigma}_1.$$

Estas relaciones de conmutación entre O_{12} y $\boldsymbol{\sigma}_1$ y $\boldsymbol{\sigma}_2$ son precisamente las mismas que las de P_{12}^{σ}, permutación que intercambia las variables de spin de los electrones 1 y 2. Luego, podemos poner

$$O_{12} = cP_{12}^{\sigma},$$

siendo c un número. La ecuación (27) demuestra que $c = \pm 1$. Para ver cuál de los dos valores de c es el adecuado, observemos que los autovalores de P_{12}^{σ} son 1, 1, 1, -1, correspondientes respectivamente a los tres estados simétricos independientes y al estado antisimétrico de las variables de spin de dos electrones, que en la notación de § 37 están representados por las tres funciones simétricas $f_\alpha(\sigma'_{z1})f_\alpha(\sigma'_{z2})$, $f_\beta(\sigma'_{z1})f_\beta(\sigma'_{z2})$, $f_\alpha(\sigma'_{z1})f_\beta(\sigma'_{z2}) + f_\beta(\sigma'_{z1})f_\alpha(\sigma'_{z2})$, y la función antisimétrica $f_\alpha(\sigma'_{z1})f_\beta(\sigma'_{z2}) - f_\beta(\sigma'_{z1})f_\alpha(\sigma'_{z2})$. Luego, la media de los autovalores de P_{12}^{σ} es $\frac{1}{2}$. Pero la media de los autovalores de $(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)$ es evidentemente igual a cero, y por tanto, la media de los autovalores de O_{12} es $\frac{1}{2}$. Luego, hemos de tomar $c = +1$, y resulta entonces

$$P_{12}^{\sigma} = \tfrac{1}{2}\{1 + (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)\}. \tag{28}$$

De esta forma toda permutación P^{σ} que consista en un simple intercambio puede expresarse como función algebraica de las $\boldsymbol{\sigma}$. Toda permutación P^{σ} puede expresarse como producto de intercambios, y en consecuencia también se podrá expresar como función algebraica de las $\boldsymbol{\sigma}$. Con ayuda de (25) podemos expresar las P^{x} como funciones algebraicas de las $\boldsymbol{\sigma}$ y eliminar las P^{σ} de la discusión. Como para los intercambios hemos de tomar el signo menos en (25), dado que el cuadrado de un intercambio es igual a uno, tenemos

$$P_{12}^{x} = -\tfrac{1}{2}\{1 + (\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2)\}. \tag{29}$$

La fórmula (29) es apta para calcular los caracteres χ' que definen los conjuntos exclusivos de estados. Por ejemplo, para las permutaciones que consisten simplemente en un intercambio, tenemos

$$\chi_{12} = \chi(P_{12}^{x}) = -\frac{1}{2}\left\{1 + \frac{2}{n(n-1)} \sum_{r<t} (\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_t)\right\}$$

Si introducimos la variable dinámica s para describir la magnitud del momento angular de spin total $\frac{1}{2} \sum_r \boldsymbol{\sigma}_r$ en unidades de $\hbar$, según la fórmula

$$s(s+1) = \left(\tfrac{1}{2} \sum_r \boldsymbol{\sigma}_r, \tfrac{1}{2} \sum_t \boldsymbol{\sigma}_t,\right)$$

que está de acuerdo con (39) de § 36, tenemos

$$2 \sum_{r<t} (\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_t) = \left(\sum_r \boldsymbol{\sigma}_r, \sum_t \boldsymbol{\sigma}_t \right) - \sum_r (\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_r)$$
$$= 4s(s+1) - 3n.$$

Luego,

$$\chi_{12} = -\tfrac{1}{2} \left\{ 1 + \frac{4s(s+1) - 3n}{n(n-1)} \right\} = -\frac{n(n-4) + 4s(s+1)}{2n(n-1)} \tag{30}$$

Por tanto, χ_{12} se puede expresar como función de las variables dinámicas s y del número de electrones n. Cualquiera de los demás χ puede calcularse de forma análoga y dicho cálculo nos llevaría a expresarlo como función únicamente de s y de n, pues no hay ninguna otra función simétrica de todas las variables dinámicas $\boldsymbol{\sigma}$ que pueda intervenir. Por tanto, existe un conjunto de valores χ' de las χ, y en consecuencia, un conjunto exclusivo de estados, para cada autovalor s' de s. Los autovalores de s son

$$\tfrac{1}{2}n, \quad \tfrac{1}{2}n - 1, \quad \tfrac{1}{2}n - 2, \quad \ldots,$$

terminando la serie con 0 o $\frac{1}{2}$.

Así vemos que cada estado estacionario de un sistema de varios electrones es un autoestado de s — magnitud del momento angular de spin total $\frac{1}{2} \sum_r \boldsymbol{\sigma}_r$ expresado en unidades de $\hbar$ — perteneciente a un autovalor determinado s'. Dado un s' cualquiera existirán $2s' + 1$ valores posibles de una componente del vector spin total en una dirección cualquiera que corresponderán a $2s' + 1$ estados estacionarios con la misma energía. Si no despreciamos las fuerzas debidas a los momentos magnéticos de spin, los $2s' + 1$ estados se desdoblarán en general en $2s' + 1$ estados con energías ligeramente diferentes, dando lugar a un multiplete de multiplicidad $2s' + 1$. Si despreciamos las fuerzas de spin, no pueden ocurrir transiciones en las que s' cambie, es decir, transiciones de una multiplicidad a otra, y si no las despreciamos sólo existirá una pequeña probabilidad de que tengan lugar.

Podemos calcular en primera aproximación los niveles de energía de un sistema de varios electrones aplicando la teoría de la sección anterior con kets $|\alpha^r\rangle$ que hagan referencia únicamente a las variables orbitales, y empleando la fórmula (23). Si únicamente consideramos las fuerzas de Coulomb entre los electrones, la energía de interacción V estará constituida por una suma de términos cada uno de los cuales hace referencia únicamente a dos electrones, de donde los elementos de matriz de V_P serán todos nulos a excepción de los que se refieran a una P^x unidad o a una P^x que consista en un simple intercambio de dos electrones. Luego, (23) se reducirá a

$$V \approx V_1 + \sum_{r<s} V_{rs} P^{\alpha}_{rs}, \tag{31}$$

siendo V_{rs} el elemento de matriz que se refiere al intercambio de los electrones r y s. Como las P^{α} tienen las mismas propiedades que las P^{σ}, cualquier función de las P^{α} tendrá los mismos autovalores que la función de las P^{σ} correspondiente y, en consecuencia, el segundo miembro de (31) tendrá los mismos autovalores que

$$V_1 + \sum_{r<s} V_{rs} P^{\sigma},$$

o

$$V_1 - \tfrac{1}{2} \sum_{r<s} V_{rs}\{1 + (\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_s)\} \tag{32}$$

según (29). Los autovalores (32) darán las correcciones de primer orden de los niveles de energía. La expresión (32) nos muestra que un modelo que suponga la existencia de una energía de acoplamiento entre los spines de los electrones que están en los estados orbitales r y s, de valor $-\frac{1}{2} V_{rs}(\boldsymbol{\sigma}_r, \boldsymbol{\sigma}_s)$ resultaría sin duda muy acertada. Esta energía de acoplamiento es mucho mayor que la de los momentos magnéticos de spin. Estos modelos atómicos se emplearon antes de conocerse la justificación que de ellos da la mecánica cuántica.

Podemos tener dos estados orbitales iguales del sistema no perturbado, es decir, que los kets $|\alpha^r\rangle$ referentes a las variables orbitales de dos electrones pueden ser iguales. Supongamos que $|\alpha^1\rangle$ y $|\alpha^2\rangle$ sean iguales. Entonces únicamente hemos de tomar los autovalores de (31) compatibles con $P^{\alpha}_{12} = 1$, o sea, aquellos autovalores de (32) compatibles con $P^{\sigma}_{12} = 1$ o $P^{\sigma} = -1$. Según (28) dicha condición da $(\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2) = -3$, de donde $(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2)^2 = 0$. Luego, la resultante de los dos spines $\boldsymbol{\sigma}_1$ y $\boldsymbol{\sigma}_2$ es nula, que puede interpretarse diciendo que los spines $\boldsymbol{\sigma}_1$ y $\boldsymbol{\sigma}_2$ son antiparalelos. Por tanto, podemos decir que dos electrones que estén en el mismo estado orbital tienen sus spines antiparalelos. No pueden haber más de dos electrones en el mismo estado orbital.

X

TEORÍA DE LA RADIACIÓN

59. *Conjuntos de bosones*

Consideremos un sistema dinámico constituido por u' partículas idénticas. Formemos una representación para una cualquiera de las partículas constituida por los kets básicos discretos $|\alpha^{(1)}\rangle$, $|\alpha^{(2)}\rangle$, $|\alpha^{(3)}\rangle$, ... A partir de ellos podemos obtener una representación simétrica del conjunto de u' partículas, como vimos en § 54, tomando como kets básicos los productos

$$|\alpha_1^a\rangle|\alpha_2^b\rangle|\alpha_3^c\rangle\ldots|\alpha_{u'}^g\rangle = |\alpha_1^a\alpha_2^b\alpha_3^c\ldots\alpha_{u'}^g\rangle \tag{1}$$

en los cuales hay un factor por cada partícula; los subíndices 1, 2, 3, ..., u' de las α son los símbolos de las partículas, y los superíndices a, b, c, ..., g representan los índices $^{(1)}$, $^{(2)}$, $^{(3)}$, ... de los kets básicos de las partículas. Si se trata de bosones, para los cuales sólo existen estados simétricos, debemos considerar únicamente los kets simétricos que pueden construirse a partir de los kets (1). Los estados correspondientes a dichos kets simétricos formarán un conjunto completo de estados para el conjunto de bosones. Podemos construir para ellos la siguiente teoría.

Introducimos el operador lineal S definido por

$$S = u'!^{-\frac{1}{2}} \sum P, \tag{2}$$

donde la suma está extendida a todas las $u'!$ permutaciones de las u' partículas. Al aplicar S a cualquier ket del conjunto se obtiene un ket simétrico. Por consiguiente, podemos llamar a S *operador de simetrización*. Según (8) de § 55, dicho operador es real. Aplicado al ket (1) da

$$u'!^{-\frac{1}{2}} \sum P|\alpha_1^a\,\alpha_2^b\,\alpha_3^c\ldots\alpha_{u'}^g\rangle = S|\alpha^a\alpha^b\alpha^c\ldots\alpha^g\rangle \tag{3}$$

En el segundo miembro se omiten los símbolos de las partículas puesto que ya no son significativos. El ket (3) representa un estado del conjunto de u' bosones definido por una distribución de los mismos entre los

diversos estados bosónicos, en la que ningún bosón particular está asociado a ningún estado particular. La distribución queda determinada cuando sabemos cuantos bosones hay en cada estado bosónico. Sean n'_1, n'_2, n'_3, ... los números de bosones que están respectivamente en los estados $\alpha^{(1)}$, $\alpha^{(2)}$, $\alpha^{(3)}$, ... Los n'_i están definidos algebraicamente por la ecuación

$$\alpha^a + \alpha^b + \alpha^c + \ldots + \alpha^g = n'_1 \alpha^{(1)} + n'_2 \alpha^{(2)} + n'_3 \alpha^{(3)} + \ldots . \tag{4}$$

La suma de las n'_i es evidentemente u'. El número de n'_i es igual al número de kets básicos $|\alpha^{(r)}\rangle$, que en la mayoría de aplicaciones de la teoría es mucho mayor que u', y así pues la mayor parte de las n'_i son nulas. Si α^a, α^b, α^c, ..., α^g son todos diferentes, es decir, si las n'_i son todas 0 o 1 el ket (3) está normalizado, puesto que en este caso los términos del primer miembro de (3) son ortogonales entre sí y cada uno contribuye con $u'!^{-1}$ al cuadrado de la longitud del ket. En cambio, si α^a, α^b, α^c, ..., α^g no son todos diferentes, los términos del primer miembro de (3) que provienen de permutaciones P, que sólo intercambian bosones en el mismo estado, serán iguales. El número de términos iguales será $n'_1!\, n'_2!\, n'_3!$ ..., y, por lo tanto, el cuadrado de la longitud del ket (3) será

$$\langle \alpha^a \alpha^b \alpha^c \ldots \alpha^g | S^2 | \alpha^a \alpha^b \alpha^c \ldots \alpha^g \rangle = n'_1!\, n'_2!\, n'_3! \ldots . \tag{5}$$

Para estudiar el estado general del conjunto podemos introducir los números n_1, n_2, n_3 ... de bosones que están respectivamente en los estados $\alpha^{(1)}$, $\alpha^{(2)}$, $\alpha^{(3)}$, ... y considerar las n_i como variables dinámicas o como observables. Sus autovalores son 0, 1, 2, ..., u'. El ket (3) es un autoket común a todas las n_i pertenecientes a los autovalores $n'_1, n'_2, n'_3, \ldots$. Los diversos kets (3) constituyen un conjunto completo para el sistema dinámico formado por u' bosones, y en consecuencia, las n_i conmutan entre sí (ver el recíproco del teorema de § 13). Además, sólo hay un ket (3) independiente que pertenezca a cualquier conjunto de autovalores $n'_1, n'_2, n'_3, \ldots$. Por lo tanto, las n_i forman un conjunto completo de observables que conmutan. Si normalizamos los kets (3) y simbolizamos los kets resultantes mediante los autovalores de las n_i a que pertenecen, es decir, si ponemos

$$(n'_1!\, n'_2!\, n'_3!\ldots)^{-\frac{1}{2}} S | \alpha^a \alpha^b \alpha^c \ldots \alpha^g \rangle = | n'_1 n'_2 n'_3 \ldots \rangle, \tag{6}$$

obtenemos un conjunto de kets $|n'_1 n'_2 n'_3 \ldots\rangle$, en los cuales todos los n'_i toman valores enteros no negativos cuya suma vale u', y que constituyen los kets básicos de una representación con las n_i diagonales.

Las n_i pueden expresarse como funciones de los observables α_1, α_2 α_3, ..., α_u, definidos por los kets básicos de los bosones individuales, mediante la ecuación

$$n_a = \sum_r \delta_{\alpha_r \alpha^a_r}, \tag{7}$$

o las ecuaciones

$$\sum_a n_a f(\alpha^a) = \sum_r f(\alpha_r) \tag{8}$$

válidas para cualquier función f.

Supongamos ahora que el número de bosones del conjunto no esté fijado, sino que sea variable. Este número es, por consiguiente, una variable dinámica u observable u, con autovalores 0, 1, 2, ..., y el ket (3) es un autoket de u perteneciente al autovalor u'. A fin de obtener un conjunto completo de kets para nuestro sistema dinámico deberemos tomar ahora todos los kets simétricos (3) para todos los valores de u'. Podemos ordenarlos del modo siguiente.

$$|\rangle, \quad |\alpha^a\rangle, \quad S|\alpha^a\alpha^b\rangle, \quad S|\alpha^a\alpha^b\alpha^c\rangle, \quad \ldots, \tag{9}$$

escribiendo en primer lugar el ket sin símbolo, que corresponde al estado en que no hay bosones, a continuación vienen los kets correspondientes a los estados con un bosón, después los correspondientes a estados con dos bosones, etc. Un estado general corresponde a un ket que es combinación lineal de los distintos kets (9). Los kets (9) son ortogonales entre sí, pues dos kets que representan estados con el mismo número de bosones son ortogonales como antes vimos, y dos kets que representan estados con distinto número de bosones son ortogonales por ser autokets de u pertenecientes a autovalores distintos. Normalizando todos los kets (9), obtenemos un conjunto de kets análogo a (6) pero para el cual no hay restricción en las n_i' (es decir, cada n_i' puede tomar todos los valores enteros no negativos), y dichos kets forman los kets básicos de una representación del sistema dinámico constituido por un número variable de bosones con las n_i diagonales.

Si no hay interacción entre los bosones y si los kets básicos $|\alpha^{(1)}\rangle$, $|\alpha^{(2)}\rangle$, ... representan estados estacionarios de un bosón, los kets (9) representarán estados estacionarios de un conjunto de bosones. El número de bosones es constante con el tiempo, pero no ha de ser necesariamente un número particular, es decir, el estado general es una superposición de estados con diversos valores de u. Si la energía de un bosón es $H(\alpha)$, la energía del conjunto será, según (8),

$$\sum_r H(\alpha_r) = \sum_a n_a H^a \tag{10}$$

designando con H^a el número $H(\alpha^a)$. Esta expresión nos da el hamiltoniano del conjunto como función de las variables dinámicas n_i.

60. *Relación entre bosones y osciladores*

En § 34 estudiábamos el oscilador armónico, sistema dinámico con un grado de libertad, que se puede describir mediante una q y p canónicas, y cuyo hamiltoniano es igual a la suma de cuadrados de q y p previamente multiplicados por coeficientes numéricos. Definimos matemáticamente un oscilador general como un sistema con un grado de libertad que puede describirse mediante una q y p canónicas, y cuyo hamiltoniano sea una serie de potencias en q y p, y que continúe siéndolo aunque el sistema sea perturbado arbitrariamente. Vamos a estudiar ahora un sistema dinámico compuesto por un conjunto de estos osciladores. Para describir cada oscilador podemos emplear en vez de q y p una variable dinámica compleja η, correspondiente a la η de § 34, y su compleja conjugada, que satisfagan la relación de conmutación (7) de § 34. Asignaremos los símbolos 1, 2, 3, ... a los diferentes osciladores, de modo que el conjunto completo de osciladores pueda describirse mediante las variables dinámicas $\eta_1, \eta_2, \eta_3, \ldots, \bar{\eta}_1, \bar{\eta}_2, \bar{\eta}_3, \ldots$, que satisfacen las relaciones de conmutación

$$\left.\begin{aligned} \eta_a \eta_b - \eta_b \eta_a &= 0, \\ \bar{\eta}_a \bar{\eta}_b - \bar{\eta}_b \bar{\eta}_a &= 0, \\ \bar{\eta}_a \eta_b - \eta_b \bar{\eta}_a &= \delta_{ab}. \end{aligned}\right\} \tag{11}$$

Pongamos

$$\eta_a \bar{\eta}_a = n_a, \tag{12}$$

de modo que

$$\bar{\eta}_a \eta_a = n_a + 1. \tag{13}$$

Las n son observables que conmutan entre sí y el desarrollo de § 34 muestra que cada una de ellas tiene como autovalores todos los enteros no negativos. Para cada oscilador a hay un ket standard de la representación de Fock $|0_a\rangle$, que es un autoket normalizado de n_a perteneciente al autovalor cero. Multiplicando entre sí todos estos kets standards obtenemos un ket standard de la representación de Fock del conjunto de osciladores

$$|0_1\rangle\, |0_2\rangle\, |0_3\rangle \ldots, \tag{14}$$

el cual es un autoket común a todas las n pertenecientes a los autovalores cero. Lo designaremos simplemente $|0\rangle$. Según (13) de § 34

$$\bar{\eta}_a |0\rangle = 0 \tag{15}$$

para cualquier a. Los desarrollos de § 34 muestran también que si $n'_1, n'_2, n'_3, \ldots$ son enteros no negativos cualesquiera,

$$\eta_1^{n'_1} \eta_2^{n'_2} \eta_3^{n'_3} \ldots |0\rangle \tag{16}$$

es un autoket común a todas las n, perteneciente a los autovalores $n'_1, n'_2, n'_3, \ldots$. Los diversos kets (16) obtenidos tomando distintas n'_i forman un conjunto completo de kets ortogonales entre sí, siendo el cuadrado de la longitud de uno de ellos, según (16) de § 34, $n'_1!\, n'_2!\, n'_3! \ldots$. Podemos ver pues, teniendo en cuenta el resultado (5), que los kets (16) tienen exactamente las mismas propiedades que los kets (9), y así igualamos cada ket (16) al ket (9) que hace referencia a los mismos valores n'_i sin introducir ninguna contradicción. Esto supone tomar

$$S|\alpha^a\alpha^b\alpha^c\ldots\alpha^g\rangle = \eta_a\,\eta_b\,\eta_c\ldots\eta_g|0\rangle. \tag{17}$$

El ket standard $|0\rangle$ resulta igual al primero de los kets (9), correspondientes a la ausencia de bosones.

La consecuencia de la ecuación (17) es la identificación de los estados de un conjunto de bosones con los estados de un conjunto de osciladores. Esto significa que *el sistema dinámico constituido por un conjunto de bosones idénticos es equivalente a un conjunto de osciladores — los dos sistemas son exactamente el mismo sistema visto desde dos puntos de vista distintos.* Hay un oscilador asociado con cada estado bosónico independiente. Éste es uno de los resultados más importantes de la mecánica cuántica, que permite realizar una unificación de las teorías ondulatoria y corpuscular de la luz.

La teoría de la sección anterior se construyó sobre un conjunto discreto de kets básicos $|\alpha^a\rangle$ para un bosón. Podríamos pasar a un conjunto discreto diferente de kets básicos $|\beta^A\rangle$, y construir sobre ellos una teoría análoga. Los kets básicos para el conjunto serían, en vez de (9)

$$|\rangle, \quad |\beta^A\rangle, \quad S|\beta^A\beta^B\rangle, \quad S|\beta^A\beta^B\beta^C\rangle, \quad \ldots. \tag{18}$$

El primero de los kets (18), que representa el estado en que no hay bosones, es el mismo que el primer ket de (9). Los kets (18), que representan estados con un bosón, son funciones lineales de los kets (9) que representan estados con un bosón, es decir

$$|\beta^A\rangle = \sum_a |\alpha^a\rangle\langle\alpha^a|\beta^A\rangle, \tag{19}$$

y en general los kets (18) referentes a la presencia de u' bosones son funciones lineales de los kets (9) referentes a la presencia de u' bosones. Así pues existirá un nuevo conjunto de variables de osciladores η_A asociadas a los nuevos estados básicos $|\beta^A\rangle$ de un bosón, y en correspondencia con (17) tendremos

$$S|\beta^A\beta^B\beta^C\ldots\rangle = \bar{\eta}_A\bar{\eta}_B\eta_C\ldots|0\rangle. \tag{20}$$

Luego un ket $\eta_A\eta_B\ldots|0\rangle$ con u' factores η_A, η_B, ... ha de ser función lineal de los kets $\eta_a\eta_b\ldots|0\rangle$ con u' factores η_a, η_b .... Por consiguiente,

cada operador lineal η_A ha de ser función lineal de las η_a. La ecuación (19) da

$$\eta_A|0\rangle = \sum_a \eta_a|0\rangle\langle\alpha^a|\beta^A\rangle$$

y de aquí

$$\eta_A = \sum_a \eta_a\langle\alpha^a|\beta^A\rangle. \tag{21}$$

Luego, *las η se transforman según la misma ley que los kets básicos de un bosón.* Las η transformadas satisfacen las mismas relaciones de conmutación con sus complejas conjugadas que las originales. Las $\bar{\eta}$ transformadas están en pie de igualdad con las originales, y por consiguiente, cuando consideramos nuestro sistema dinámico como un conjunto de osciladores, los diferentes grados de libertad no tienen un significado invariante.

Las $\bar{\eta}$ se transforman según la misma ley que los bras básicos de un bosón, y por lo tanto, según la misma ley que los números $\langle\alpha^a|x\rangle$ que forman el representante de un estado x. Esta similitud da lugar a que a menudo se diga que las $\bar{\eta}$ vienen dadas por un proceso de *segunda cuantificación* aplicado a $\langle\alpha^a|x\rangle$, significando con ello que, tras haber construido una teoría cuántica para una sola partícula e introducido los números $\langle\alpha^a|x\rangle$ que representan un estado de la partícula, pueden introducirse operadores lineales correspondientes que satisfacen las relaciones de conmutación correctas (11) con sus compleos conjugados, obteniéndose así la base matemática apropiada para tratar un conjunto de dichas partículas, siempre que se trate de bosones. Para fermiones hay un proceso correspondiente, que estudiaremos en § 65.

Como un conjunto de bosones es equivalente a un conjunto de osciladores, ha de ser posible expresar cualquier función simétrica de las variables del bosón en función de las variables η y $\bar{\eta}$ del oscilador. La ecuación (10) nos proporciona un ejemplo de ello con $\eta_a\bar{\eta}_a$ sustituido por n_a. Veamos qué ocurre en general. Consideremos en primer lugar una función de las variables del bosón de la forma

$$U_T = \sum_r U_r, \tag{22}$$

siendo cada U_r función solamente de las variables dinámicas del bosón r-ésimo, de donde su representante $\langle\alpha_r^a|U_r|\alpha_r^b\rangle$ únicamente se refiere a los kets básicos $|\alpha_r^a\rangle$ del bosón r-ésimo. Para que U_T sea simétrica, este representante ha de ser el mismo para todo r, luego sólo puede depender de los dos autovalores simbolizados por a y b. Para abreviar podemos, pues, escribirlo

$$\langle\alpha_r^a|U_r|\alpha_r^b\rangle = \langle\alpha^a|U|\alpha^b\rangle = \langle a|U|b\rangle \tag{23}$$

Tenemos

$$U_r|\alpha_1^{x_1}\,\alpha_2^{x_2}\ldots\rangle = \sum_a |\alpha_1^{x_1}\,\alpha_2^{x_2}\ldots\alpha_r^a\ \ldots\rangle\langle a|U|x_r\rangle. \tag{24}$$

Sumando esta ecuación para todos los valores de r y aplicando el operador de simetrización S a ambos miembros, obtenemos

$$SU_T|\alpha_1^{x_1}\,\alpha_2^{x_2}\ldots\rangle = \sum_r \sum_a S|\alpha_1^{x_1}\,\alpha_2^{x_2}\ldots\alpha_r^a\ \ldots\rangle\langle a|U|x_r\rangle. \tag{25}$$

Como U_T es simétrica podemos reemplazar SU_T *por* U_TS y sustituir entonces los kets simétricos de (25) por sus expresiones dadas por (17). De este modo obtenemos

$$\begin{aligned} U_T\,\eta_{x_1}\eta_{x_2}\ldots|0\rangle &= \sum_a \sum_r \eta_a\ \eta_{x_r}^{-1}\eta_{x_1}\eta_{x_2}\ldots|0\rangle\langle a|U|x_r\rangle \\ &= \sum_{ab} \eta_a \sum_r \eta_{x_r}^{-1}\eta_{x_1}\eta_{x_2}\ldots\ |0\rangle\delta_{bx_r}\langle a|U|b\rangle, \end{aligned} \tag{26}$$

$\eta_{x_r}^{-1}$ significa que el factor η_{x_r} ha de ser suprimido. A partir de (15) y de las relaciones de conmutación (11) se tiene

$$\bar{\eta}_b\eta_{x_1}\eta_{x_2}\ldots|0\rangle = \sum_r \eta_{x_r}^{-1}\eta_{x_1}\eta_{x_2}\ldots|0\rangle\delta_{bx_r}, \tag{27}$$

(obsérvese que $\bar{\eta}_b$ actúa como el operador de derivación parcial $\partial/\partial\eta_b$), de donde (26) da

$$U_T\,\eta_{x_1}\,\eta_{x_2}\ldots|0\rangle = \sum_{ab} \eta_a\,\bar{\eta}_b\,\eta_{x_1}\,\eta_{x_2}\ldots|0\rangle\langle a|U|b\rangle. \tag{28}$$

Los kets $\bar{\eta}_{x_1}\,\eta_{x_2}\ldots|0\rangle$ forman un conjunto completo, luego, a partir de (28) podemos inferir la ecuación entre operadores

$$U_T = \sum_{ab} \eta_a\langle a|U|b\rangle\bar{\eta}_b. \tag{29}$$

que nos da U_T en función de las variables η y $\bar{\eta}$ y de los elementos de matriz $\langle a|U|b\rangle$.

Tomemos ahora una función simétrica de las variables del bosón que consista en una suma de términos referidos cada uno a dos bosones.

$$V_T = \sum_{r,s\neq r} V_{rs}. \tag{30}$$

No necesitamos suponer $V_{rs} = V_{sr}$. De acuerdo con (33), a fin de abreviar, podemos escribir los elementos de matriz de V_{rs} en la forma

$$\langle\alpha_r^a\,\alpha_s^b|V_{rs}|\alpha_r^c\,\alpha_s^d\rangle = \langle ab|V|cd\rangle \tag{31}$$

Procediendo como antes, obtenemos el resultado correspondiente a (25)

$$SV_T|\alpha_1^{x_1}\,\alpha_2^{x_2}\ldots\rangle = \sum_{r,s\neq r} \sum_{ab} S|\ \alpha_1^{x_1}\alpha_2^{x_2}\ldots\alpha_r^a\ldots\alpha_s^b\ldots\rangle\langle ab|V|x_rx_s\rangle \tag{32}$$

y el que corresponde a (26)

$$V_T \eta_{x_1} \eta_{x_2} \ldots |0\rangle = \sum_{abcd} \eta_a \eta_b \sum_{r,s\neq r} \eta^{-1}_{x_r} \eta^{-1}_{x_s} \eta_{x_1} \eta_{x_2} \ldots |0\rangle \delta_{cx_r} \delta_{dx_s} \langle ab|V|cd\rangle. \quad (33)$$

Por generalización de (27) podemos deducir

$$\bar{\eta}_c \bar{\eta}_d \eta_{x_1} \eta_{x_2} \ldots |0\rangle = \sum_{r,s\neq r} \eta^{-1}_{x_r} \eta^{-1}_{x_s} \eta_{x_1} \eta_{x_2} \ldots |0\rangle \delta_{cx_r} \delta_{dx_s}, \quad (34)$$

de donde (33) da

$$V_T \eta_{x_1} \eta_{x_2} \ldots |0\rangle = \sum_{abcd} \eta_a \eta_b \bar{\eta}_c \bar{\eta}_d \eta_{x_1} \eta_{x_2} \ldots |0\rangle \langle ab|V|cd\rangle,$$

que nos lleva a la ecuación entre operadores

$$V_T = \sum_{abcd} \eta_a \eta_b \langle ab|V|cd\rangle \bar{\eta}_c \bar{\eta}_d. \quad (35)$$

El método puede generalizarse fácilmente para expresar cualquier función simétrica de las variables del bosón en función de las η y las $\bar{\eta}$.

La teoría precedente puede generalizarse con facilidad para aplicarla a un conjunto de bosones en interacción con algún otro sistema dinámico como, por ejemplo, el átomo. Tendremos que introducir un conjunto de kets básicos $|\zeta'\rangle$ para el átomo solo. A partir de ellos podemos obtener un conjunto de kets básicos para el sistema total, formado por el átomo y los bosones, multiplicando cada ket $|\zeta'\rangle$ por cada uno de los kets (9). Podemos escribir estos kets

$$|\zeta'\rangle, \quad |\zeta'\alpha^a\rangle, \quad S|\zeta'\alpha^a\alpha^b\rangle, \quad S|\zeta'\alpha^a\alpha^b\alpha^c\rangle, \quad \ldots. \quad (36)$$

Consideremos el sistema como compuesto por el átomo en interacción con un conjunto de osciladores, pudiéndose describir entonces en función de las variables del átomo y de las variables η_a, $\bar{\eta}_a$ de los osciladores. Volviendo a usar el ket standard $|0\rangle$ del conjunto de osciladores, tenemos

$$S|\zeta'\alpha^a\alpha^b\alpha^c \ldots\rangle = \eta_a \eta_b \eta_c \ldots |0\rangle |\zeta'\rangle, \quad (37)$$

en correspondencia con (17), como ecuación que expresa los kets básicos (36) en función de las variables del oscilador.

Cualquier función de las variables del átomo y de las variables de los bosones que sea simétrica respecto a todos los bosones puede expresarse como función de las variables del átomo y de las η y $\bar{\eta}$. Consideremos en primer lugar una función U_T de la forma (22) con U_r función solamente de las variables del átomo y de las variables del bosón r-ésimo, cuyo representante será pues $\langle\zeta'\alpha^a_r|U_r|\zeta''\alpha^b_r\rangle$. Este representante ha de ser independiente de r para que U_T pueda ser simétrica respecto a todos los bosones, luego, podemos escribirlo $\langle\zeta'\alpha^a|U|\zeta''\alpha^b\rangle$. Definamos ahora $\langle a|U|b\rangle$

como la función de las variables atómicas cuyo representante sea $\langle\zeta'\alpha^a|U|\zeta''\alpha^b\rangle$; tenemos pues

$$\langle\zeta'\alpha_r^a|U_r|\zeta''\alpha_r^b\rangle = \langle\zeta'\alpha^a|U|\zeta''\alpha^b\rangle = \langle\zeta'|\langle a|U|b\rangle|\zeta''\rangle, \tag{38}$$

en correspondencia con (23). Podemos recurrir ahora a las ecuaciones (24) - (28) y aplicarlas multiplicando previamente ambos miembros por $|\zeta'\rangle$ a la derecha, de donde se deduce que la fórmula (29) continúa siendo válida. Podemos tratar análogamente la función simétrica V_T de la forma (30), con V_{rs} función solamente de las variables atómicas y de las variables de los bosones r-ésimo y s-ésimo. Definiendo $\langle ab|V|cd\rangle$ como la función de las variables atómicas cuyo representante es

$$\langle\zeta'\alpha_r^a\,\alpha_s^b|V_{rs}|\zeta''\alpha_r^c\,\alpha_s^d\rangle,$$

la fórmula (35) continúa siendo válida.

61. *Emisión y absorción de bosones*

Supongamos que los osciladores de la sección anterior son armónicos y que no hay interacción entre ellos. La energía del oscilador a-ésimo es, según (5) de § 34,

$$H_a = \hbar\omega_a\,\eta_a\,\bar{\eta}_a + \tfrac{1}{2}\hbar\omega_a.$$

Despreciaremos el término constante $\frac{1}{2}\hbar\omega_a$ que es la energía del oscilador en su nivel más bajo — el llamado "punto cero de energía" —. Esta omisión no tiene ninguna consecuencia dinámica, como queda explicado al principio de § 30, e implica simplemente una nueva definición de H_a. La energía total de todos los osciladores es ahora

$$H_T = \sum_a H_a = \sum_a \hbar\omega_a\,\eta_a\,\bar{\eta}_a = \sum_a \hbar\omega_a\,n_a \tag{39}$$

con la ayuda de (12). Es de la misma forma que (10) con $\hbar\omega_a$ en lugar de H^n. Así pues *un conjunto de osciladores armónicos es equivalente a un conjunto de bosones en estados estacionarios sin interacción entre ellos. Si un oscilador del conjunto está en su estado cuántico n', hay n' bosones en el estado bosónico asociado.*

En general, el hamiltoniano del conjunto de osciladores será una serie de potencias en las variables η, $\bar{\eta}$, o sea,

$$H_T = H_P + \sum_a (U_a\,\eta_a + \bar{U}_a\,\bar{\eta}_a) + \sum_{ab}(U_{ab}\,\eta_a\,\bar{\eta}_b + V_{ab}\,\eta_a\,\eta_b + \bar{V}_{ab}\,\bar{\eta}_a\,\bar{\eta}_b) + \dots \tag{40}$$

donde H_P, U_a, U_{ab}, V_{ab} son números, siendo H_P real, y $U_{ab} = \bar{U}_{ba}$. Si el conjunto de osciladores está en interacción con un átomo, como teníamos

al final de la sección precedente, el hamiltoniano total será también de la forma (40), con H_P, U_a, U_{ab}, V_{ab} funciones de las variables atómicas siendo en particular H_P el hamiltoniano del átomo. El estudio general de este sistema dinámico sería más bien complicado y para las aplicaciones prácticas se supone que los términos

$$H_P + \sum_a U_{aa} \dot{\eta}_a \bar{\eta}_a \tag{41}$$

son grandes en comparación con los otros y constituyen por sí solos un sistema no perturbado; los restantes términos se consideran como una perturbación que produce transiciones en el sistema no perturbado, de acuerdo con la teoría de § 44. Si además U_{aa} es independiente de las variables atómicas, el sistema sin perturbar cuyo hamiltoniano es (41) está formado simplemente por un átomo de hamiltoniano H_P y un conjunto de bosones en estados estacionarios con el hamiltoniano en la forma (39) sin interacción.

Consideremos qué clases de transiciones producen los diversos términos de perturbación de (40). Sea un estado estacionario del sistema sin perturbar para el cual el átomo está en un estado estacionario ζ', y los bosones permanecen en los estados bosónicos estacionarios, a, b, c, ... Este estado estacionario para el sistema sin perturbar corresponde al ket

$$\dot{\eta}_a \dot{\eta}_b \dot{\eta}_c \ldots |0\rangle|\zeta'\rangle, \tag{42}$$

análoga a (37). Si se multiplica este ket por el término $U_x \dot{\eta}_x$ de (40), el resultado es una combinación lineal de kets del tipo

$$\eta_x \dot{\eta}_a \dot{\eta}_b \dot{\eta}_c \ldots |0\rangle|\zeta''\rangle, \tag{43}$$

en la cual ζ'' es cualquier estado estacionario del átomo. El ket (43) representa un estado con un bosón más que el ket (42); este bosón extra está en el estado x. Luego el término de perturbación $U_x \dot{\eta}_x$ origina transiciones en las que se emite un bosón en el estado x y el átomo pasa a un estado arbitrario. Si aplicamos el término $\bar{U}_x \bar{\eta}_x$ de (40) al ket (42), el resultado es cero a no ser que (42) contenga un factor $\dot{\eta}_x$ en cuyo caso será una combinación lineal de kets de la forma

$$\eta_x^{-1} \dot{\eta}_a \dot{\eta}_b \dot{\eta}_c \ldots |0\rangle|\zeta''\rangle,$$

con un bosón menos en el estado x. Luego el término de perturbación $\bar{U}_x \bar{\eta}_x$ origina transiciones en las que se absorbe un bosón del estado x, y el átomo pasa también a un estado arbitrario. Análogamente puede verse que un término de perturbación $U_{xy} \dot{\eta}_x \bar{\eta}_y$ $(x \neq y)$ origina procesos en los que se absorbe un bosón en el estado y y se emite uno en el estado x, o, lo que es lo mismo desde el punto de vista físico, un bosón hace una

transición del estado y al estado x. Este tipo de procesos sería producido por un término como U_T de (22) y (29) en la energía de perturbación, cuyos elementos diagonales $\langle a|U|a\rangle$ fueran nulos. Por otra parte los términos de perturbación $V_{xy}\,\eta_x\,\eta_y$, $\bar{V}_{xy}\,\bar{\eta}_x\,\bar{\eta}_y$ originan procesos en los que son emitidos o absorbidos dos bosones, y así sucesivamente para términos más complicados. En cualquiera de estos procesos de emisión y absorción el átomo puede pasar a un estado arbitrario.

Determinemos la dependencia de la probabilidad de que ocurra cada uno de estos procesos de transición respecto a la distribución original de bosones en los diversos estados bosónicos. Según §§ 44 y 46 la probabilidad de transición es siempre proporcional al cuadrado del módulo del elemento de matriz de la energía de perturbación, que hace referencia a los dos estados en cuestión. Así pues la probabilidad de que un bosón sea emitido en el estado x y que el átomo pase simultáneamente del estado ζ' al estado ζ'' es proporcional a

$$|\langle\zeta''|\langle n'_1 n'_2 \ldots (n'_x + 1)\ldots|U_x \eta_x|n'_1 n'_2 \ldots n'_x \ldots\rangle|\zeta'\rangle|^2 \tag{44}$$

las n'_i son el número de bosones presentes inicialmente en los diversos estados bosónicos. De (6) y (17), teniendo en cuenta (4), resulta

$$|n'_1 n'_2 n'_3 \ldots\rangle = (n'_1!n'_2!n'_3!\ldots)^{\frac{1}{2}}\eta_1^{n'_1}\,\eta_2^{n'_2}\,\eta_3^{n'_3}\ldots|0\rangle, \tag{45}$$

y así pues

$$\bar{\eta}_x|n'_1 n'_2 \ldots n'_x \ldots\rangle = (n'_x + 1)^{\frac{1}{2}}|n'_1 n'_2 \ldots (n'_x + 1)\ldots\rangle. \tag{46}$$

Luego, (44) es igual a

$$(n'_x + 1)|\langle\zeta''|U_x|\zeta'\rangle|^2, \tag{47}$$

es decir, que *la probabilidad de una transición en la que se emite un bosón en el estado x es proporcional al número de bosones que estaban anteriormente en el estado x, más uno.*

La probabilidad de que sea absorbido un bosón en el estado x y que el átomo pase simultáneamente del estado ζ' al estado ζ'' es proporcional a

$$|\langle\zeta''|\langle n'_1 n'_2 \ldots (n'_x - 1)\ldots|\bar{U}_x \bar{\eta}_x|n'_1 n'_2 \ldots n'_x \ldots\rangle|\zeta'\rangle|^2 \tag{48}$$

donde las n'_i, igual que antes, son el número de bosones presentes inicialmente en los diversos estados bosónicos. De (45)

$$\bar{\eta}_x|n'_1 n'_2 \ldots n'_x \ldots\rangle = n'^{\frac{1}{2}}_x|n'_1 n'_2 \ldots (n'_x - 1)\ldots\rangle, \tag{49}$$

luego, (48) es igual a

$$n'_x|\langle\zeta''|\bar{U}_x|\zeta'\rangle|^2 \tag{50}$$

Es decir, la probabilidad de que tenga lugar una transición en la que se absorba un bosón en el estado x es proporcional al número de bosones originalmente en el estado x.

Pueden aplicarse métodos análogos para procesos más complicados y ver que la probabilidad de que un bosón haga una transición del estado y al estado x ($x \neq y$) es proporcional a $n'_y(n'_x + 1)$. Más en general, la probabilidad de que sean absorbidos bosones en los estados x, y, ... y emitidos en los estados a, b, ... es proporcional a

$$n'_x n'_y \ldots (n'_a + 1)(n'_b + 1) \ldots, \tag{51}$$

siendo las n'_y en cada caso el número de bosones presentes inicialmente. Estos resultados valen para procesos de transición directa y para procesos de transición que se realizan a través de uno o más estados intermedios, de acuerdo con la interpretación dada al final de § 44.

62. *Aplicación a fotones*

Puesto que los fotones son bosones, puede aplicárseles la teoría precedente. Los estados estacionarios de un fotón son los autoestados de su momento. Para cada autoestado del momento existen dos estados de polarización independientes, para los cuales pueden elegirse dos estados de polarización lineal en direcciones perpendiculares. Las variables dinámicas necesarias para describir los estados estacionarios son en este caso el vector momento $\mathbf{p}$ y una variable de polarización $\mathbf{l}$ consistente en un vector unidad perpendicular a $\mathbf{p}$. Las variables $\mathbf{p}$ y $\mathbf{l}$ ocupan el lugar de las α anteriores. Los autovalores de $\mathbf{p}$ son todos los números desde $-\infty$ a ∞ para cada una de las tres componentes cartesianas de $\mathbf{p}$, y para cada autovalor p' de $\mathbf{p}$, $\mathbf{l}$ tiene solamente dos autovalores que corresponden a dos vectores escogidos de manera arbitraria, perpendiculares a $\mathbf{p}'$ y perpendiculares entre sí. Puesto que los autovalores de $\mathbf{p}$ forman un intervalo continuo, existirá un conjunto continuo de estados estacionarios que constituyen los kets básicos continuos $|\mathbf{p}'\mathbf{l}'\rangle$. En cambio, la teoría precedente se construyó para un conjunto discreto de kets básicos $|\alpha'\rangle$ de un bosón. Podemos recurrir a dos formalismos para eliminar esta discrepancia.

El primero consiste en reemplazar la distribución continua tridimensional de autovalores de $\mathbf{p}$ por un gran número de puntos discretos situados muy cerca unos de otros y formando una nube esparcida sobre todo el espacio $\mathbf{p}$ tridimensional. Sea $s_{\mathbf{p}'}$ la densidad de la nube (número de puntos por unidad de volumen) en las proximidades de cualquier punto $\mathbf{p}'$. Por consiguiente, $s_{\mathbf{p}'}$ ha de ser grande y positivo, pero aparte de eso puede ser una función arbitraria de $\mathbf{p}'$. Las integrales sobre el espacio $\mathbf{p}$ pueden reemplazarse por sumas extendidas a la nube de puntos, de acuerdo con la fórmula

$$\int\int\int f(\mathbf{p}')\, dp'_x\, dp'_y\, dp'_z = \sum f(\mathbf{p}') s^{-1}_{\mathbf{p}'}, \tag{52}$$

la cual nos proporciona la base para el paso de los valores de $\mathbf{p}'$ continuos a los discretos y viceversa. Cualquier problema puede plantearse en términos de los valores discretos de $\mathbf{p}'$, para los cuales puede usarse la teoría de §§ 59 - 61, y los resultados pueden transformarse de nuevo para referirlos a los valores continuos de $\mathbf{p}'$. La densidad arbitraria $s_{\mathbf{p}'}$ desaparece entonces de los resultados.

El segundo formalismo consiste en modificar las ecuaciones de la teoría de §§ 59 - 61 de modo que se apliquen al caso de un conjunto de kets básicos $|\alpha'\rangle$ continuo, reemplazando sumas por integrales y el símbolo δ de las relaciones de conmutación (11) por las funciones δ en lo que concierne a las variables con autovalores continuos. Cada uno de estos formalismos tiene sus ventajas y sus inconvenientes. El primero suele ser más conveniente para la discusión física, y el segundo para el desarrollo matemático. Desarrollaremos los dos y emplearemos uno u otro según nos convenga en cada momento.

El hamiltoniano que describe un conjunto de fotones en interacción con un átomo tendrá la forma general (40), con coeficientes H_P, U_a, U_{ab}, V_{ab} que contienen las variables atómicas. Este hamiltoniano puede escribirse

$$H_T = H_P + H_Q + H_R, \tag{53}$$

donde H_P es la energía del átomo solo, H_R la energía del conjunto de fotones aislado

$$H_R = \sum_{\mathbf{p}'\lambda'} n_{\mathbf{p}'\lambda'} h\nu_{\mathbf{p}'}, \tag{54}$$

— $\nu_{\mathbf{p}'}$ es la frecuencia de un fotón de momento $\mathbf{p}'$ —, y H_Q es la energía de interacción que puede calcularse por analogía con la teoría clásica como veremos en la sección siguiente. El sistema global puede tratarse por un método de perturbación como el de la sección precedente, con H_P y H_R en el lugar de la energía (41) del sistema no perturbado y H_Q como energía de perturbación que origina las transiciones con emisión y absorción de fotones y pasos simultáneos del átomo de un estado estacionario a otro.

Vimos en la sección anterior que la probabilidad de que tenga lugar un proceso de absorción es proporcional al número de bosones que estaban inicialmente en el estado en que es absorbido el bosón. Podemos deducir de ello que la probabilidad de que un fotón procedente de un haz de radiación que incide sobre un átomo sea absorbido es proporcional a la intensidad del haz. Vimos también que la probabilidad de un proceso de emisión es proporcional al número de bosones presentes anteriormente en el estado dado más uno. Para interpretar este resultado hemos de hacer un estudio cuidadoso de las relaciones implicadas al reemplazar el conjunto continuo de estados fotónicos por un conjunto discreto.

Despreciemos por ahora la variable de polarización l. Sea $|\mathbf{p}'\mathrm{D}\rangle$ el ket normalizado correspondiente al estado fotónico discreto $\mathbf{p}'$. Entonces según (22) de § 16

$$\sum_{\mathbf{p}'} |\mathbf{p}'\mathrm{D}\rangle\langle\mathbf{p}'\mathrm{D}| = 1,$$

que da, según (52),

$$\int |\mathbf{p}'\mathrm{D}\rangle\langle\mathbf{p}'\mathrm{D}|s_{\mathbf{p}'}\, d^3p' = 1, \tag{55}$$

habiendo escrito para abreviar d^3p' en lugar de $dp'_x dp'_y dp'_z$. Si $|\mathbf{p}'\rangle$ es el ket básico correspondiente al estado continuo $\mathbf{p}'$, según (24) de § 16, tenemos

$$\int |\mathbf{p}'\rangle\langle\mathbf{p}'|d^3p' = 1,$$

que muestra, comparando con (55) que

$$|\mathbf{p}'\rangle = |\mathbf{p}'\mathrm{D}\rangle s^{\frac{1}{2}}_{\mathbf{p}'} \tag{56}$$

La relación entre $|\mathbf{p}'\rangle$ y $|\mathbf{p}'\mathrm{D}\rangle$ es análoga a la que existe entre los kets básicos al cambiar la función peso de la representación; véase (38) de § 16.

Si tenemos $n'_{\mathbf{p}'}$ fotones en cada estado fotónico discreto $\mathbf{p}'$, la densidad de Gibbs ρ para el conjunto de fotones es, según (68) de § 33,

$$\begin{aligned}\rho &= \sum_{\mathbf{p}'} |\mathbf{p}'\mathrm{D}\rangle n'_{\mathbf{p}'}\langle\mathbf{p}'\mathrm{D}| = \int |\mathbf{p}'\mathrm{D}\rangle n'_{\mathbf{p}'}\, \langle\mathbf{p}'\mathrm{D}|s_{\mathbf{p}'}\, d^3p' \\ &= \int |\mathbf{p}'\rangle n'_{\mathbf{p}'}\langle\mathbf{p}'|d^3p'\end{aligned} \tag{57}$$

con la ayuda de (56). El número de fotones por unidad de volumen en el entorno de un punto cualquiera es, por consiguiente, $\langle\mathbf{x}'|\rho|\mathbf{x}'\rangle$, según (73) de § 33.

Según (57), al introducir el valor de la función de transformación $\langle\mathbf{x}'|\mathbf{p}'\rangle$ dada por (54) de § 23, da como resultado

$$\begin{aligned}\langle\mathbf{x}'|\rho|\mathbf{x}'\rangle &= \int \langle\mathbf{x}'|\mathbf{p}'\rangle n'_{\mathbf{p}'}\langle\mathbf{p}'|\mathbf{x}'\rangle\, d^3p' \\ &= \int h^{-3} n'_{\mathbf{p}'} d^3p'\end{aligned} \tag{58}$$

La ecuación (58) expresa el número de fotones por unidad de volumen como una integral sobre el espacio de los momentos, luego, el integrando de (58) puede interpretarse como el número de fotones por unidad de volumen del espacio de fases. Llegamos de esta manera al resultado que *el número de fotones por unidad de volumen del espacio de fases es igual a h^{-3} veces el número de fotones por estado discreto,* o dicho de otro modo, *una celda de volumen h^3 en el espacio de fases es equivalente a un estado discreto.* Este resultado es general y válido para cualquier clase de partículas. Si se tiene en cuenta la variable de polarización de los

fotones, el resultado vale para cada uno de los dos estados de polarización independientes.

La magnitud del momento de un fotón de frecuencia ν es $h\nu/c$, luego, el elemento del espacio de momentos es

$$dp_x dp_y dp_z = h^3 c^{-3} \nu^2 d\nu d\omega,$$

siendo $d\omega$ un elemento de ángulo sólido en la dirección del vector $\mathbf{p}$. Así pues, una distribución de fotones que contenga $n'_{\mathbf{p}}$ fotones por estado discreto, o lo que es equivalente, una distribución de $h^{-3} n'_{\mathbf{p}} d^3p d^3x$ fotones por elemento de volumen d^3x y un elemento del espacio de momentos d^3p, es igual a una distribución de $n'_{\mathbf{p}} c^{-3} \nu^2 d\nu d\omega d^3x$ fotones por elemento de volumen d^3x, por intervalo de frecuencia $d\nu$ y por elemento de ángulo sólido $d\omega$ en la dirección $\mathbf{p}$ del movimiento. Esto corresponde a una densidad de energía $n'_{\mathbf{p}} hc^{-3}\nu^3$ por unidad de ángulo sólido y por unidad de intervalo de frecuencia, o a una intensidad por unidad de intervalo de frecuencia (que es la energía que atraviesa la unidad de área por unidad de tiempo y por unidad de intervalo de frecuencia) de valor

$$I_\nu = n'_{\mathbf{p}} h\nu^3/c^2. \tag{59}$$

El hecho de que la probabilidad de emisión de un fotón sea proporcional a $n'_{\mathbf{p}1} + 1$, siendo $n'_{\mathbf{p}1}$ el número de fotones que inicialmente están en el estado discreto dado, puede interpretarse ahora como que la probabilidad es proporcional a $I_{\nu 1} + h\nu^3/c^2$, donde $I_{\nu 1}$ es la intensidad de la radiación incidente por unidad de intervalo de frecuencias en el entorno de la frecuencia del fotón emitido y con su misma polarización 1. Por consiguiente, incluso cuando no incide radiación sigue existiendo una cierta emisión, pero la emisión resulta incrementada o *estimulada* por la radiación incidente en la misma dirección y con la misma frecuencia y polarización de la radiación emitida. La teoría de la radiación aquí desarrollada completa de este modo la teoría deficiente de § 45, justificando tanto la emisión estimulada como la espontánea. La relación que da para los dos tipos de emisión, $I_{\nu 1} : h\nu^3/c^2$, está en concordancia con la que proporciona la teoría de Einstein del equilibrio estadístico, mencionada en § 45.

La probabilidad de que en un experimento de dispersión pase un fotón del estado $\mathbf{p}'1'$ al estado $\mathbf{p}''1''$ es proporcional a $n_{\mathbf{p}'1'}(n_{\mathbf{p}''1''} + 1)$, donde las n son el número de fotones que estaban inicialmente en los estados discretos dados. Podemos interpretar este resultado diciendo que la probabilidad es proporcional a

$$I_{\nu'1'}(I_{\nu''1''} + h\nu''^3/c^2). \tag{60}$$

De manera semejante para un proceso de radiación más general en el que se emiten y absorben varios fotones, la probabilidad es proporcional al

factor $I_{\nu 1}$ para cada fotón absorbido y al factor $I_{\nu 1} + h\nu^3/c^2$ para cada fotón emitido. Luego, el proceso es estimulado por la radiación incidente en la misma dirección y con la misma frecuencia y polarización que los fotones emitidos.

63. *La energía de interacción de los fotones con un átomo*

Vamos a determinar ahora la energía de interacción entre un átomo y un conjunto de fotones, es decir, la H_Q de la ecuación (53), por analogía con la expresión clásica de la energía de interacción entre un átomo y un campo de radiación. Para simplificar supondremos que el átomo está formado por un solo electrón que se mueve en un campo de fuerzas electrostático. El campo de radiación puede describirse mediante un potencial escalar y un potencial vector. Estos potenciales no están unívocamente determinados y pueden escogerse de manera que el potencial escalar desaparezca. El campo queda entonces completamente determinado por el potencial vector A_x, A_y, A_z o $\mathbf{A}$. El cambio en el hamiltoniano que describe el átomo producido por el campo es, según hemos visto al principio de § 41,

$$H_Q = \frac{1}{2m}\left\{\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - \mathbf{p}^2\right\} = \frac{e}{mc}(\mathbf{p}, \mathbf{A}) + \frac{e^2}{2mc^2}\mathbf{A}^2. \tag{61}$$

Ésta es la energía clásica de interacción. El $\mathbf{A}$ de esta expresión tendría que ser el valor del potencial vector en el punto en que está situado el electrón en el instante considerado. Sin embargo, si tomamos para $\mathbf{A}$ el valor del potencial vector en algún punto fijo del átomo, como por ejemplo el núcleo, resulta una aproximación suficientemente buena, siempre que se trate de radiaciones cuya longitud de onda sea grande en comparación con las dimensiones del átomo.

Consideremos primero clásicamente el campo de radiación e ignoremos su interacción con el átomo. Según la teoría de Maxwell, el potencial vector satisface las ecuaciones

$$\Box \mathbf{A} = 0, \qquad \operatorname{div} \mathbf{A} = 0, \tag{62}$$

$\Box$ simboliza $\partial^2/c^2\partial t^2 - \partial^2/\partial x^2 - \partial^2/\partial y^2 - \partial^2/\partial z^2$. La primera ecuación nos dice que $\mathbf{A}$ puede descomponerse en componentes de Fourier de la forma

$$\mathbf{A} = \int \{\mathbf{A}_\mathbf{k}\, e^{-i(\mathbf{kx})+2\pi i\nu_\mathbf{k} t} + \bar{\mathbf{A}}_\mathbf{k}\, e^{i(\mathbf{kx})-2\pi i\nu_\mathbf{k} t}\}\, d^3k, \tag{63}$$

Cada componente de Fourier representa un tren de ondas que avanza con la velocidad de la luz, dado por un vector $\mathbf{k}$ cuya dirección da la dirección

de movimiento de las ondas y cuya magnitud $|\mathbf{k}|$ está relacionada con su frecuencia $\nu_{\mathbf{k}}$ por

$$2\pi\nu_{\mathbf{k}} = c|\mathbf{k}|. \tag{64}$$

El vector $\mathbf{k}$ es precisamente el momento del fotón que la teoría cuántica asociaría con estas ondas dividido por $\hbar$. Para cada valor de $\mathbf{k}$ tenemos una amplitud $\mathbf{A}_{\mathbf{k}}$, que, en general, es un vector complejo, y la integral de (63) se extiende sobre todo el espacio tridimensional de las $\mathbf{k}$. La segunda de las ecuaciones (62) da

$$(\mathbf{k}, \mathbf{A}_{\mathbf{k}}) = 0, \tag{65}$$

que nos dice que para cada valor de $\mathbf{k}$, $\mathbf{A}_{\mathbf{k}}$ es perpendicular a $\mathbf{k}$. Esto expresa que las ondas son transversales. $\mathbf{A}_{\mathbf{k}}$ está determinado por sus dos componentes en dos direcciones perpendiculares entre sí y a su vez perpendiculares a $\mathbf{k}$; estas dos componentes corresponden a dos estados de polarización lineal independientes.

La energía total de radiación viene dada por la integral de volumen

$$H_R = (8\pi)^{-1}\int (\mathcal{E}^2 + \mathcal{H}^2)\, d^3x \tag{66}$$

extendida a todo el espacio; el campo eléctrico $\mathcal{E}$ y el campo magnético $\mathcal{H}$ vienen dados por

$$\mathcal{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \qquad \mathcal{H} = \operatorname{rot} \mathbf{A}. \tag{67}$$

Usando fórmulas conocidas del análisis vectorial, resulta

$$\begin{aligned}\operatorname{div}[\mathbf{A} \times \mathcal{H}] &= (\mathcal{H}, \operatorname{rot} \mathbf{A}) - (\mathbf{A}, \operatorname{rot} \mathcal{H}) = \mathcal{H}^2 - (\mathbf{A}, \operatorname{rot}\operatorname{rot} \mathbf{A}) \\ &= \mathcal{H}^2 + (\mathbf{A}, \nabla^2\mathbf{A})\end{aligned}$$

con la ayuda de la segunda ecuación (62). De este modo, despreciando un término que puede transformarse en una integral de superficie extendida a la superficie del infinito, (66) se convierte en

$$H_R = (8\pi)^{-1}\int\left\{\frac{1}{c^2}\left(\frac{\partial \mathbf{A}}{\partial t}, \frac{\partial \mathbf{A}}{\partial t}\right) - (\mathbf{A}, \nabla^2\mathbf{A})\right\} d^3x. \tag{68}$$

Sustituyendo $\mathbf{A}$ en esta expresión por su valor (63) podemos obtener la energía de radiación en función de las amplitudes de Fourier $\mathbf{A}_{\mathbf{k}}$. La energía de la radiación es constante (puesto que despreciamos la interacción de la radiación con el átomo), luego para calcularla podemos tomar $t = 0$. Esto significa tomar

$$\mathbf{A} = \int (\mathbf{A}_{\mathbf{k}} + \bar{\mathbf{A}}_{-\mathbf{k}})e^{-i(\mathbf{kx})}\, d^3k, \tag{69}$$

$$\nabla^2\mathbf{A} = -\int \mathbf{k}^2 (\mathbf{A}_{\mathbf{k}} + \bar{\mathbf{A}}_{-\mathbf{k}})e^{-i(\mathbf{kx})}\, d^3k,$$

$$\partial\mathbf{A}/\partial t = ic \int |\mathbf{k}|(\mathbf{A}_{\mathbf{k}} - \bar{\mathbf{A}}_{-\mathbf{k}})e^{-i(\mathbf{kx})}\, d^3k. \tag{70}$$

Introduciendo estas expresiones en (68), obtenemos, con la ayuda de la fórmula (49) de § 23

$$\begin{aligned} H_R &= (8\pi)^{-1} \int\int\int \{\mathbf{k}'^2(\mathbf{A}_\mathbf{k} + \bar{\mathbf{A}}_{-\mathbf{k}}, \mathbf{A}_{\mathbf{k}'} + \bar{\mathbf{A}}_{-\mathbf{k}'}) - \\ &\quad - |\mathbf{k}|\,|\mathbf{k}'|(\mathbf{A}_\mathbf{k} - \bar{\mathbf{A}}_{-\mathbf{k}}, \mathbf{A}_{\mathbf{k}'} - \bar{\mathbf{A}}_{-\mathbf{k}'})\} e^{-i(\mathbf{k}\mathbf{x})} e^{-i(\mathbf{k}'\mathbf{x})}\, d^3k d^3k' d^3x \\ &= \pi^2 \int\int \{\mathbf{k}'^2(\mathbf{A}_\mathbf{k} + \bar{\mathbf{A}}_{-\mathbf{k}}, \mathbf{A}_{\mathbf{k}'} + \bar{\mathbf{A}}_{-\mathbf{k}'}) - \\ &\quad - |\mathbf{k}|\,|\mathbf{k}'|(\mathbf{A}_\mathbf{k} - \bar{\mathbf{A}}_{-\mathbf{k}}, \mathbf{A}_{\mathbf{k}'} - \bar{\mathbf{A}}_{-\mathbf{k}'})\} \delta(\mathbf{k} + \mathbf{k}')\, d^3k d^3k', \end{aligned}$$

siendo $\delta(\mathbf{k} + \mathbf{k}')$ el producto de tres factores correspondientes a las tres componentes de $\mathbf{k}$. De ahí

$$\begin{aligned} H_R &= \pi^2 \int \mathbf{k}^2 \{(\mathbf{A}_\mathbf{k} + \bar{\mathbf{A}}_{-\mathbf{k}}, \mathbf{A}_{-\mathbf{k}} + \bar{\mathbf{A}}_\mathbf{k}) - (\mathbf{A}_\mathbf{k} - \bar{\mathbf{A}}_{-\mathbf{k}}, \mathbf{A}_{-\mathbf{k}} - \bar{\mathbf{A}}_\mathbf{k})\}\, d^3k \\ &= 2\pi^2 \int \mathbf{k}^2 \{(\mathbf{A}_\mathbf{k}, \bar{\mathbf{A}}_\mathbf{k}) + (\mathbf{A}_{-\mathbf{k}}, \bar{\mathbf{A}}_{-\mathbf{k}})\}\, d^3k \\ &= 4\pi^2 \int \mathbf{k}^2 (\mathbf{A}_\mathbf{k}, \bar{\mathbf{A}}_\mathbf{k})\, d^3k. \end{aligned} \tag{71}$$

Podemos reemplazar la distribución de valores de $\mathbf{k}$ continua por una nube de valores discretos de $\mathbf{k}$, como hicimos con los valores de $\mathbf{p}$ en la sección anterior. La integral (71) se transforma entonces, según la fórmula (52) en la suma

$$H_R = 4\pi^2 \sum_\mathbf{k} \mathbf{k}^2 (\mathbf{A}_\mathbf{k}, \bar{\mathbf{A}}_\mathbf{k}) s_\mathbf{k}^{-1},$$

siendo $s_\mathbf{k}$ la densidad de valores discretos de $\mathbf{k}$. Podemos escribirlo también en la forma

$$H_R = 4\pi^2 \sum_{\mathbf{k}\mathbf{l}} \mathbf{k}^2 A_{\mathbf{k}\mathbf{l}} \bar{A}_{\mathbf{k}\mathbf{l}}\, s_\mathbf{k}^{-1}, \tag{72}$$

siendo $A_{\mathbf{k}\mathbf{l}}$ la componente de $\mathbf{A}_\mathbf{k}$ en una dirección $\mathbf{l}$ perpendicular a $\mathbf{k}$ y estando referida la sumación respecto a $\mathbf{l}$ a dos direcciones $\mathbf{l}$ perpendiculares entre sí. Es decir, en (72) hay un término para cada estado estacionario de un fotón independiente.

Los valores $\mathcal{E}$ y $\mathcal{H}$ del campo en un punto cualquiera pueden tomarse como variables dinámicas. Las cantidades

$$A_{\mathbf{k}\mathbf{l}t} = A_{\mathbf{k}\mathbf{l}}\, e^{2\pi i \nu_\mathbf{k} t}, \qquad \bar{A}_{\mathbf{k}\mathbf{l}t} = \bar{A}_{\mathbf{k}\mathbf{l}}\, e^{2\pi i \nu_\mathbf{k} t},$$

son entonces variables dinámicas en el instante t, puesto que están ligadas a $\mathcal{E}$ y $\mathcal{H}$ en diversos puntos $\mathbf{x}$ y en el instante t por ecuaciones que no contienen la t, como se deduce de (63) y (67). $A_{\mathbf{k}\mathbf{l}}$ es constante, luego, $A_{\mathbf{k}\mathbf{l}t}$ varía con t según una ley armónica simple. Por lo tanto, $A_{\mathbf{k}\mathbf{l}t}$ hace

el papel de la η_t de un oscilador armónico, definido por (3) de § 34, donde la ω del oscilador es $2\pi\nu_{\mathbf{k}}$. Podemos tomar cada $A_{\mathbf{kl}t}$ proporcional a la η_t de un oscilador armónico, y entonces el campo de radiación se convierte en un conjunto de osciladores armónicos.

Pasemos a la teoría cuántica y tomemos $A_{\mathbf{kl}t}$, $\bar{A}_{\mathbf{kl}t}$ como variables dinámicas en la imagen de Heisenberg. La expresión (72) de la energía puede conservarse, si bien ahora el orden en que aparecen los factores $A_{\mathbf{kl}}$ y $\bar{A}_{\mathbf{kl}}$ no puede modificarse si no queremos que aparezca un punto cero de energía. Las $A_{\mathbf{kl}t}$ continúan variando con el tiempo de acuerdo con la ley $e^{i\omega t}$, e igual que antes pueden tomarse proporcionales a las η_t de osciladores armónicos. El factor de proporcionalidad puede obtenerse igualando (72) a la expresión (39) de la energía, sustituyendo en ella el símbolo a por los dos símbolos $\mathbf{k}$ y $\mathbf{l}$, y $\hbar\omega_a$ por $h\nu_{\mathbf{k}}$. Esto nos da

$$4\pi^2 \sum_{\mathbf{kl}} \mathbf{k}^2 A_{\mathbf{kl}t}\bar{A}_{\mathbf{kl}t}s_{\mathbf{k}}^{-1} = \sum_{\mathbf{kl}} h\nu_{\mathbf{k}}\, \eta_{\mathbf{kl}t}\, \bar{\eta}_{\mathbf{kl}t},$$

Introducimos el subíndice t para indicar que trabajamos con variables dinámicas de Heisenberg (como hemos de hacer cada vez que pasamos ecuaciones de la teoría clásica a la teoría cuántica). De ahí, usando (64),

$$4\pi^2 A_{\mathbf{kl}t} = ch^{\frac{1}{2}}\nu_{\mathbf{k}}^{-\frac{1}{2}}\eta_{\mathbf{lk}t}s_{\mathbf{k}}^{\frac{1}{2}}, \tag{73}$$

despreciando un factor de fase arbitrario sin importancia. De este modo se introducen las variables dinámicas de Heisenberg $\eta_{\mathbf{kl}t}$, que describen el campo de radiación como un conjunto de osciladores. Las relaciones de conmutación entre $\eta_{\mathbf{kl}t}$ y $\bar{\eta}_{\mathbf{kl}t}$ son conocidas y están dadas por (11), luego, la ecuación (73) fija las relaciones de conmutación entre $A_{\mathbf{kl}t}$ y $\bar{A}_{\mathbf{kl}t}$. En consecuencia, fija también las relaciones de conmutación entre los potenciales $\mathbf{A}$ y los valores del campo $\boldsymbol{\mathcal{E}}$ y $\boldsymbol{\mathcal{H}}$ en los diversos puntos $\mathbf{x}$ y en el instante $\mathbf{t}$. (Incidentalmente las relaciones de conmutación entre $A_{\mathbf{kl}}$ y $\bar{A}_{\mathbf{kl}}$ no dependen del tiempo, y, por consiguiente, la relación de conmutación entre dos potenciales o dos cantidades que describen el campo en dos tiempos distintos tampoco varía con el tiempo.)

Asimismo podemos usar (73) cuando queremos tener en cuenta la interacción entre el campo de radiación y el átomo. Esta afirmación implica suponer que la interacción no afecta a las relaciones de conmutación entre los potenciales y las cantidades que describen el campo en un instante dado. La interacción hace que las $\eta_{\mathbf{kl}t}$ dejen de variar de acuerdo con la ley armónica simple y que los osciladores dejen de ser armónicos. Por consiguiente, podría afectar a la relación de conmutación entre dos potenciales o dos cantidades que describen el campo para dos instantes distintos.

Podemos introducir ahora la energía de interacción (61) en la teoría cuántica sustituyendo $\mathbf{p}$ por $\mathbf{p}_t$ para indicar que es una variable dinámica

de Heisenberg. Tomando el núcleo atómico como origen, al sustituir (63) con $\mathbf{x} = 0$ en (61), obtenemos

$$H_{Qt} = \frac{e}{mc}\int (\mathbf{p}_t, \mathbf{A}_{\mathbf{k}t} + \bar{\mathbf{A}}_{\mathbf{k}t})\, d^3k + \\ + \frac{e^2}{2mc^2}\iint (\mathbf{A}_{\mathbf{k}t} + \bar{\mathbf{A}}_{\mathbf{k}t}, \mathbf{A}_{\mathbf{k}'t} + \bar{\mathbf{A}}_{\mathbf{k}'t})\, d^3k d^3k'$$

$$= \frac{e}{mc}\sum_{\mathbf{k}} (\mathbf{p}_t, \mathbf{A}_{\mathbf{k}t} + \bar{\mathbf{A}}_{\mathbf{k}t}) s_{\mathbf{k}}^{-1} + \frac{e^2}{2mc^2}\sum_{\mathbf{k}\mathbf{k}'} (\mathbf{A}_{\mathbf{k}t} + \bar{\mathbf{A}}_{\mathbf{k}t}, \mathbf{A}_{\mathbf{k}'t} + \bar{\mathbf{A}}_{\mathbf{k}'t})\, s_{\mathbf{k}}^{-1} s_{\mathbf{k}'}^{-1}$$

si pasamos de valores continuos de $\mathbf{k}$ a valores discretos. Luego,

$$H_{Qt} = \frac{e}{mc}\sum_{\mathbf{k}\mathbf{l}} p_{\mathbf{l}t}(A_{\mathbf{k}\mathbf{l}t} + \bar{A}_{\mathbf{k}\mathbf{l}t}) s_{\mathbf{k}}^{-1} + \\ + \frac{e^2}{2mc^2}\sum_{\mathbf{k}\mathbf{k}'\mathbf{l}\mathbf{l}'} (A_{\mathbf{k}\mathbf{l}t} + \bar{A}_{\mathbf{k}\mathbf{l}t})(A_{\mathbf{k}'\mathbf{l}'t} + \bar{A}_{\mathbf{k}'\mathbf{l}'t})(\mathbf{l}\mathbf{l}') s_{\mathbf{k}}^{-1} s_{\mathbf{k}'}^{-1},$$

donde $p_{\mathbf{l}t}$ es la componente de $\mathbf{p}_t$ en la dirección $\mathbf{l}$. Con la ayuda de (73) podemos expresar H_{Qt} en función de $\eta_{\mathbf{k}\mathbf{l}t}$ y $\bar{\eta}_{\mathbf{k}\mathbf{l}t}$, y eliminar entonces el subíndice t (lo que significa pasar a las variables dinámicas de Schrödinger, obteniéndose finalmente

$$H_Q = \frac{eh^{\frac{1}{2}}}{4\pi^2 m}\sum_{\mathbf{k}\mathbf{l}} p_{\mathbf{l}} \nu_{\mathbf{k}}^{-\frac{1}{2}} (\eta_{\mathbf{k}\mathbf{l}} + \bar{\eta}_{\mathbf{k}\mathbf{l}}) s_{\mathbf{k}}^{-\frac{1}{2}} + \\ + \frac{e^2 h}{32\pi^4 m}\sum_{\mathbf{k}\mathbf{k}'\mathbf{l}\mathbf{l}'} \nu_{\mathbf{k}}^{-\frac{1}{2}} \nu_{\mathbf{k}'}^{-\frac{1}{2}} (\eta_{\mathbf{k}\mathbf{l}} + \bar{\eta}_{\mathbf{k}\mathbf{l}})(\eta_{\mathbf{k}'\mathbf{l}'} + \bar{\eta}_{\mathbf{k}'\mathbf{l}'})(\mathbf{l}\mathbf{l}')\, s_{\mathbf{k}}^{-\frac{1}{2}} s_{\mathbf{k}'}^{-\frac{1}{2}}. \qquad (74)$$

Para el modelo de átomo que empleamos, la energía de interacción aparece como una función lineal más una función cuadrática en las η y las $\bar{\eta}$. Los términos lineales originan procesos de emisión y absorción, y los cuadráticos procesos de dispersión y procesos en los que se absorben o emiten dos fotones simultáneamente. El orden de los factores η y $\bar{\eta}$ en los términos cuadráticos no queda determinado por el método empleado en la teoría clásica, pero no tiene importancia puesto que un cambio del orden no modifica a H_Q más que en una constante.

El elemento de matriz de H_Q relativo a la emisión de un fotón en el estado discreto $\mathbf{k}\mathbf{l}$, o lo que es lo mismo, en el estado discreto $\mathbf{p}'\mathbf{l}$ como también puede designarse, acompañada del salto del átomo del estado α^0 al estado α', es

$$\langle \mathbf{p}'\mathbf{l}\alpha' | H_Q | \alpha^0 \rangle = \frac{eh^{\frac{1}{2}}}{4\pi^2 m \nu'^{\frac{1}{2}}} \langle \alpha' | p_{\mathbf{l}} | \alpha^0 \rangle s_{\mathbf{k}}^{-\frac{1}{2}} = \frac{e}{mh(2\pi\nu')^{\frac{1}{2}}} \langle \alpha' | p_{\mathbf{l}} | \alpha^0 \rangle s_{\mathbf{p}}^{-\frac{1}{2}}$$

puesto que $s_{\mathbf{k}} = s_{\mathbf{p}}\hbar^3$. El $p_{\mathbf{l}}$ que figura aquí, y que se refiere al momento del electrón es, por supuesto, totalmente distinto de las otras letras $\mathbf{p}$, que

se refieren al momento del fotón emitido. Para evitar confusiones reemplazaremos el momento **p** del electrón por $m\dot{x}$, ya que, para el átomo sin perturbar, estas dos variables dinámicas son una misma. Pasando a estados continuos del fotón mediante la imaginaria conjugada de la ecuación (56), obtenemos

$$\langle \mathbf{p'l}\alpha'|H_Q|\alpha^0\rangle = \frac{e}{h(2\pi\nu')^{\frac{1}{2}}}\langle\alpha'|\dot{x}_1|\alpha_0\rangle. \tag{75}$$

Análogamente, el elemento de matriz de H_Q relativo a la absorción de un fotón en el estado continuo $\mathbf{p}^0\mathbf{l}$, pasando el átomo simultáneamente del estado α^0 al estado α' es

$$\langle\alpha'|H_Q|\mathbf{p}^0\mathbf{l}\alpha^0\rangle = \frac{e}{h(2\pi\nu^0)^{\frac{1}{2}}}\langle\alpha'|\dot{x}_1|\alpha^0\rangle, \tag{76}$$

y el elemento de matriz relativo a la dispersión de un fotón desde el estado continuo $\mathbf{p}^0\mathbf{l}^0$ al estado continuo $\mathbf{p'l'}$ con el átomo saltando del estado α^0 al estado α' es

$$\langle\mathbf{p'l'}\alpha'|H_Q|\mathbf{p}^0\mathbf{l}^0\alpha^0\rangle = \frac{e^2}{2\pi h^2 m\nu^{0\frac{1}{2}}\nu'^{\frac{1}{2}}}(\mathbf{l'l}^0)\delta_{\alpha'\alpha^0} \tag{77}$$

para el cual hay dos términos de (74) que dan contribución no nula. En la próxima sección emplearemos estos elementos de matriz. Los elementos de matriz relativos a la absorción o emisión simultánea de dos fotones pueden obtenerse de la misma manera, pero conducen a efectos físicos demasiado pequeños para ser de importancia práctica.

64. *Emisión, absorción y dispersión de la radiación*

Podemos determinar ahora directamente los coeficientes de emisión, absorción y dispersión de radiaciones sustituyendo en las fórmulas del capítulo VIII los valores de los elementos de matriz dados por (75), (76) y (77).

Para determinar la probabilidad de emisión podemos usar la fórmula (56) de § 53. Nos dice que para un átomo en un estado α^0 la probabilidad por unidad de tiempo y por unidad de ángulo sólido de que emita espontáneamente un fotón y pase a un estado α' de menor energía es

$$\frac{4\pi^2}{h}\frac{WP}{c^2}\left|\frac{e}{h}\frac{1}{(2\pi\nu)^{\frac{1}{2}}}\langle\alpha'|\dot{x}_1|\alpha^0\rangle\right|^2 \tag{78}$$

Ahora bien, la energía y el momento de un fotón de frecuencia ν son

$$W = h\nu, \qquad P = h\nu/c.$$

Además, de la ley de Heisenberg (20) de § 29,

$$\langle\alpha'|\dot{x}_1|\alpha^0\rangle = -2\pi i\nu(\alpha^0\alpha')\langle\alpha'|x_1|\alpha^0\rangle,$$

donde $\nu(\alpha^0\alpha')$ es la frecuencia asociada a las transiciones del estado α^0 al estado α', que en nuestro caso es precisamente la frecuencia ν de la radiación emitida. Estos resultados sustituidos en (78) reducen el coeficiente de emisión a

$$\frac{(2\pi\nu)^3}{hc^3}|\langle\alpha'|ex_1|\alpha^0\rangle|^2. \tag{79}$$

Para obtener la cantidad de energía emitida en la unidad de tiempo por unidad de ángulo sólido con una polarización dada tenemos que multiplicarlo por $h\nu$. Con ello, el coeficiente total de emisión de energía en todas direcciones será

$$\frac{4}{3}\frac{(2\pi\nu)^4}{c^3}|\langle\alpha'|e\mathbf{x}|\alpha^0\rangle|^2, \tag{80}$$

que está de acuerdo con la expresión (34) de § 45 y justifica la hipótesis de Heisenberg para la interpretación de sus elementos de matriz.

De la misma manera el coeficiente de absorción, dado por la fórmula (59) de § 53, se convierte para fotones en

$$\frac{4\pi^2h^2W}{c^2P}\left|\frac{e}{h}\frac{1}{(2\pi\nu)^{\frac{1}{2}}}\langle\alpha'|\dot{x}_1|\alpha^0\rangle\right|^2 = \frac{8\pi^3\nu}{c}|\langle\alpha'|ex_1|\alpha^0\rangle|^2.$$

Este coeficiente de absorción se refiere a un haz incidente que contiene un fotón por unidad de intervalo de energía que atraviesa la unidad de área en la unidad de tiempo. Si en lugar de por unidad de intervalo de energía tomamos por unidad de intervalo de frecuencia, como se suele hacer cuando se trata de radiaciones, el coeficiente de absorción resulta ser

$$\frac{8\pi^3\nu}{hc}|\alpha'|ex_1|\alpha^0\rangle|^2.$$

Este resultado es el mismo que (32) de § 45, si en él sustituimos E_ν por la energía $h\nu$ de un solo fotón. De este modo *la teoría elemental de § 45, en la que se considera el campo de radiación como una perturbación externa, da el valor correcto para el coeficiente de absorción.*

Esta concordancia entre la teoría elemental y la presente teoría podría deducirse de argumentos generales. Las dos teorías difieren solamente en que las cantidades que describen el campo conmutan entre sí en la teoría elemental y, en cambio, en la presente teoría satisfacen unas relaciones de conmutación definidas, diferencia que resulta sin importancia para campos fuertes. Por lo tanto, las dos teorías han de dar la misma emisión

y absorción cuando intervienen campos fuertes. Como las dos teorías dan el coeficiente de absorción proporcional a la intensidad del haz incidente, la concordancia ha de mantenerse también para campos débiles en el caso de la absorción. De la misma manera la parte de emisión estimulada de la presente teoría ha de coincidir con la emisión en la teoría elemental.

Consideremos ahora la dispersión. El coeficiente de dispersión directa viene dado por la fórmula (38) de § 50. Esta dispersión de fotones no irá acompañada por ningún cambio en el estado del átomo, debido a la presencia del factor $\delta_{\alpha'\alpha^0}$ en la expresión del elemento de matriz (77). Luego, la energía final W' del fotón será igual a su energía inicial W^0. El coeficiente de dispersión queda reducido a

$$e^4/m^2c^4 \cdot (\mathbf{l}'\mathbf{l}^0)^2.$$

Es el mismo que da la mecánica clásica para la dispersión de radiación por un electrón libre. Vemos, pues, que la dispersión directa de radiación por un electrón en un átomo es independiente del átomo y viene dada correctamente por la teoría clásica. Recordemos que este resultado sólo es válido cuando la longitud de onda de la radiación es grande en comparación con las dimensiones del átomo.

La dispersión directa es un concepto matemático y no puede separarse experimentalmente de la dispersión total, dada por la fórmula (44) de § 51. Veamos cuál es esta dispersión total en el caso de fotones. Hemos de ir con cuidado al aplicar la fórmula (44) de § 51. La sumación $\sum_k$ en esta fórmula representa la contribución a la dispersión de transiciones dobles, consistentes en transiciones primeramente del estado inicial al estado k y luego del estado k al estado final. La primera transición puede ser una absorción del fotón incidente y la segunda una emisión del fotón dispersado, pero también es posible que tenga primero lugar la emisión y después la absorción. La naturaleza general del método usado para deducir la fórmula (44) de § 51 pone de manifiesto que cuando se aplica a fotones han de incluirse los dos tipos de transición doble en la sumación $\sum_k$ aunque en la deducción dada en § 51 solamente aparezcan las del primer tipo debido a que allí no se tuvo en cuenta la posibilidad de que la partícula fuera creada o aniquilada.

Empleamos los superíndices cero, prima y segunda para referirnos respectivamente a los estados inicial, final e intermedio del átomo, y cero y prima para referirnos respectivamente a los fotones absorbidos y emitidos. Para la doble transición de absorción seguida de emisión tendremos que tomar entonces como elementos de matriz

$$\langle k|V|\mathbf{p}^0\alpha^0\rangle, \qquad \langle \mathbf{p}'\alpha'|V|k\rangle$$

en la fórmula (44) de § 51 los siguientes:

$$\langle k|V|\mathbf{p}^0\alpha^0\rangle = \langle \alpha''|H_Q|\mathbf{p}^0\mathbf{l}^0\alpha^0\rangle, \qquad \langle \mathbf{p}'\alpha'|V|k\rangle = \langle \mathbf{p}'\mathbf{l}'\alpha'|H_Q|\alpha''\rangle.$$

También hemos de tomar

$$E' - E_k = h\nu^0 + H_P(\alpha^0) - H_P(\alpha'') = h[\nu^0 - \nu(\alpha''\alpha^0)],$$

donde

$$h\nu(\alpha''\alpha^0) = H_P(\alpha'') - H_P(\alpha^0).$$

Análogamente, para la doble transición de emisión seguida de absorción hemos de tomar

$$\langle k|V|\mathbf{p}^0\alpha^0\rangle = \langle \mathbf{p}'\mathbf{l}'\alpha''|H_Q|\alpha^0\rangle, \qquad \langle \mathbf{p}'\alpha'|V|k\rangle = \langle \alpha'|H_Q|\mathbf{p}^0\mathbf{l}^0\alpha''\rangle$$

y

$$E' - E_k = h\nu^0 + H_P(\alpha^0) - H_P(\alpha'') - h\nu^0 - h\nu' = -h[\nu' + \nu(\alpha''\alpha^0)],$$

existiendo ahora en el estado intermedio dos fotones, de frecuencias ν^0 y ν'. Sustituyendo los valores de los elementos de matriz dados por (75), (76) y (77) en (44) de § 51 obtenemos para el coeficiente de dispersión

$$\frac{e^4}{h^2c^4}\,\frac{\nu'}{\nu^0}\left|\frac{h}{m}(\mathbf{l}'\mathbf{l}^0)\,\delta_{\alpha'\alpha^0} + \right.$$
$$\left. + \sum_{\alpha''}\left\{\frac{\langle\alpha'|\dot{x}_{1'}|\alpha''\rangle\langle\alpha''|\dot{x}_{1^0}|\alpha^0\rangle}{\nu^0 - \nu(\alpha''\alpha^0)} - \frac{\langle\alpha'|\dot{x}_{1^0}|\alpha''\rangle\langle\alpha''|\dot{x}_{1'}|\alpha^0\rangle}{\nu' + \nu(\alpha''\alpha^0)}\right\}\right|^2. \tag{81}$$

Poniendo (81) en función de x en vez de $\dot{x}$, obtenemos

$$\frac{(2\pi e)^4}{h^2c^4}\,\frac{\nu'}{\nu^0}\left|\frac{\hbar}{2\pi m}(\mathbf{l}'\mathbf{l}^0)\,\delta_{\alpha'\alpha^0} - \sum_{\alpha''}\nu(\alpha'\alpha'')\nu(\alpha''\alpha^0)\left\{\frac{\langle\alpha'|x_{1'}|\alpha''\rangle\langle\alpha''|x_{1^0}|\alpha^0\rangle}{\nu^0 - \nu(\alpha''\alpha^0)} - \right.\right.$$
$$\left.\left. - \frac{\langle\alpha'|x_{1^0}|\alpha''\rangle\langle\alpha''|x_{1'}|\alpha^0\rangle}{\nu' + \nu(\alpha''\alpha^0)}\right\}\right|^2. \tag{82}$$

Podemos simplificar (82) mediante las condiciones cuánticas. Tenemos

$$x_{1'}x_{1^0} - x_{1^0}x_{1'} = 0,$$

que da

$$\sum_{\alpha''}\{\langle\alpha'|x_{1'}|\alpha''\rangle\langle\alpha''|x_{1^0}|\alpha^0\rangle - \{\langle\alpha'|x_{1^0}|\alpha''\rangle\langle\alpha''|x_{1'}|\alpha^0\rangle\} = 0, \tag{83}$$

y también

$$x_{1'}\dot{x}_{1^0} - \dot{x}_{1^0}x_{1'} = 1/m\cdot(x_{1'}p_{1^0} - p_{1^0}x_{1'}) = i\hbar/m\cdot(\mathbf{l}'\mathbf{l}^0),$$

que da

$$\sum_{\alpha''}\{\langle\alpha'|x_{1'}|\alpha''\rangle\cdot\nu(\alpha''\alpha^0)\langle\alpha''|x_{1^0}|\alpha^0\rangle - \nu(\alpha'\alpha'')\langle\alpha'|x_{1^0}|\alpha''\rangle\cdot\langle\alpha''|x_{1'}|\alpha^0\rangle\}$$
$$= \frac{1}{2\pi i}\,\frac{i\hbar}{m}(\mathbf{l}'\mathbf{l}^0)\delta_{\alpha'\alpha^0} = \frac{\hbar}{2\pi m}(\mathbf{l}'\mathbf{l}^0)\delta_{\alpha'\alpha^0}. \tag{84}$$

Multiplicando (83) por ν' y sumándola a (84) obtenemos

$$\sum_{\alpha''}\{\langle\alpha'|x_{1'}|\alpha''\rangle\langle\alpha''|x_{1^0}|\alpha^0\rangle[\nu' + \nu(\alpha''\alpha^0)] - \langle\alpha'|x_{1^0}|\alpha''\rangle\langle\alpha''|x_{1'}|\alpha^0\rangle[\nu' + \nu(\alpha'\alpha'')]\}$$
$$= \hbar/2\pi m\cdot(\mathbf{l}'\mathbf{l}^0)\,\delta_{\alpha'\alpha^0}.$$

Si sustituimos en (82) $\hbar/2\pi m \cdot (\mathbf{1}^0)\,\delta_{\alpha'\alpha^0}$ por esta expresión, obtenemos después de una simplificación directa haciendo uso de relaciones idénticas entre las ν,

$$\frac{(2\pi e)^4}{h^2c^4}\,\nu^0\nu'^3\left|\sum_{\alpha''}\left\{\frac{\langle\alpha'|x_{1'}|\alpha''\rangle\langle\alpha''|x_{1^0}|\alpha^0\rangle}{\nu^0-\nu(\alpha''\alpha^0)}-\frac{\langle\alpha'|x_{1^0}|\alpha''\rangle\langle\alpha''|x_{1'}|\alpha^0\rangle}{\nu'+\nu(\alpha''\alpha^0)}\right\}\right|^2. \tag{85}$$

que es el coeficiente de dispersión en la forma de sección eficaz sobre la que ha de incidir el fotón por unidad de ángulo sólido de dispersión. Se conoce como *fórmula de dispersión de Kramer-Heisenberg*, y fue obtenida primeramente por estos autores mediante analogías con la teoría clásica de dispersión.

El hecho de que los diversos términos de (82) puedan combinarse para dar el resultado (85) justifica la suposición hecha al deducir la fórmula (44) de § 51, de que los elementos de matriz $\langle\mathbf{p}'\alpha'|V|\mathbf{p}''\alpha''\rangle$ de la energía de interacción son de segundo orden de magnitud comparados con los $\langle\mathbf{p}'\alpha'|V|\mathbf{k}\rangle$ en todo promedio temporal cuando las partículas son fotones.

65. *Conjuntos de fermiones*

Un conjunto de fermiones puede estudiarse mediante un método análogo al usado en §§ 59 y 60 para bosones. Con los kets (1) podemos usar el operador de antisimetrización A definido por

$$A = u'!^{-\frac{1}{2}}\sum \pm P, \tag{2'}$$

donde la suma está extendida a todas las permutaciones P, y con el signo $+$ o $-$ según sea P par o impar. Aplicado al ket (1) da

$$u'!^{-\frac{1}{2}}\sum \pm P|\alpha_1^a\alpha_2^b\alpha_3^c\ldots\alpha_{u'}^g\rangle = A|\alpha^a\alpha^b\alpha^c\ldots\alpha^g\rangle, \tag{3'}$$

ket correspondiente a un estado de un conjunto de u' fermiones. El ket (3′) está normalizado si los kets $|\alpha^a\rangle$, $|\alpha^b\rangle$, ... de los fermiones individuales son todos diferentes, en caso contrario es cero. A este respecto el ket (3′) es más simple que el ket (3). En cambio (3′) es más complicado que (3) por el hecho de que (3′) depende del orden en que se presentan α^a, α^b, α^c, ..., y cambia de signo cuando se aplica a este orden una permutación impar.

Podemos introducir igual que antes los números de fermiones n_1, n_2, n_3, ... en los estados $\alpha^{(1)}$, $\alpha^{(2)}$, $\alpha^{(3)}$, ... y considerarlos como variables dinámicas u observables. Cada uno de ellos tiene por autovalores solamente 0 y 1. Forman un conjunto completo de observables que conmutan para el conjunto de fermiones. Los kets básicos de una representación con las n_i diagonales pueden tomarse ligados a los kets (3′) por la ecuación

$$A|\alpha^a\alpha^b\alpha^c\ldots\alpha^g\rangle = \pm\,|\,n_1'n_2'n_3'\ldots\rangle \tag{6'}$$

correspondiente a (6); las n'_i están relacionadas con las variables α^a, α^b α^c, ... por la ecuación (4.) En (6′) es necesario poner el signo $\pm$, puesto que para unas n'_i dadas los estados ocupados α^a, α^b, α^c, ... quedan fijados, pero no su orden y, por tanto, el signo del primer miembro de (6′) tampoco está fijado. Para establecer una regla que determine el signo de (6′) hemos de componer arbitrariamente todos los estados α de un fermión en un orden dado. Las α del primer miembro de (6′) son un subconjunto de todas las α y el orden establecido para todas las α fijará un orden para el subconjunto. Fijemos ahora la siguiente regla: pondremos signo más en (6′), si las α del primer miembro pueden ponerse en el orden previamente establecido, mediante una permutación par, y el signo menos, si para ello es necesario aplicar una permutación impar. Debido a la complejidad de esta regla la representación con los kets básicos $|n'_1 n'_2 n'_3 \ldots\rangle$ no es muy útil.

Si el número de fermiones en el conjunto es variable, podemos formar un conjunto completo de kets

$$|\rangle, \quad |\alpha^a\rangle, \quad A|\alpha^a\alpha^b\rangle, \quad A|\alpha^a\alpha^b\alpha^c\rangle, \quad \ldots, \tag{9′}$$

correspondiente a (9). Un ket general puede expresarse ahora como combinación lineal de los diversos kets (9′).

Para continuar con este desarrollo introducimos un conjunto de operadores lineales η, $\bar{\eta}$ (un par de ellos η_a, $\bar{\eta}_a$ para cada estado fermiónico α^a), que verifiquen las relaciones de conmutación

$$\left.\begin{aligned} \eta_a\,\eta_b + \eta_b\,\eta_a &= 0, \\ \bar{\eta}_a\,\bar{\eta}_b + \bar{\eta}_b\,\bar{\eta}_a &= 0, \\ \bar{\eta}_a\,\eta_b + \eta_b\,\bar{\eta}_a &= \delta_{ab}. \end{aligned}\right\} \tag{11′}$$

Estas relaciones son iguales a las (11) pero con un signo + en vez de — en el primer miembro. Para $a \neq b$, η_a y $\bar{\eta}_a$ anticonmutan con η_b y $\bar{\eta}_b$, mientras que para $b = a$ dan

$$\eta_a^2 = 0, \qquad \bar{\eta}_a^2 = 0, \qquad \bar{\eta}_a\,\eta_a + \eta_a\,\bar{\eta}_a = 1. \tag{11″}$$

Para comprobar que las relaciones (11′) no se contradicen veamos cómo pueden construirse operadores η, $\bar{\eta}$ que verifiquen las condiciones (11′). Para cada estado α^a tomemos un conjunto de operadores lineales σ_{xa}, σ_{ya}, σ_{za} como los σ_x, σ_y, σ_z introducidos en § 37 para describir el spin de un electrón, tales que σ_{xa}, σ_{ya}, σ_{za} conmutan con σ_{xb}, σ_{yb}, σ_{zb} para $b \neq a$. Tomemos también un conjunto de operadores lineales ζ_a independientes, uno para cada estado α^a, que anticonmuten todos entre sí, cuyos cuadrados sean la unidad y que conmuten con todas las variables σ. Entonces, poniendo

$$\eta_a = \tfrac{1}{2}\zeta_a(\sigma_{xa} - i\sigma_{ya}), \quad \bar{\eta}_a = \tfrac{1}{2}\zeta_a(\sigma_{xa} + i\sigma_{ya}),$$

se satisfacen todas las condiciones (11′).

De (11″)

$$(\eta_a \bar{\eta}_a)^2 = \eta_a \bar{\eta}_a \eta_a \bar{\eta}_a = \eta_a(1 - \eta_a \bar{\eta}_a)\bar{\eta}_a = \eta_a \bar{\eta}_a.$$

Que es una ecuación algebraica en $\eta_a \bar{\eta}_a$ de la cual deducimos que $\eta_a \bar{\eta}_a$ es un observable cuyos autovalores son 0 y 1. Además $\eta_a \bar{\eta}_a$ conmuta con $\eta_b \bar{\eta}_b$ para $b \neq a$. Estos resultados nos permiten poner

$$\eta_a \bar{\eta}_a = n_a, \tag{12′}$$

igual que en (12). De (11″) obtenemos ahora

$$\bar{\eta}_a \eta_a = 1 - n_a, \tag{13′}$$

ecuación correspondiente a (13).

Llamemos $|0\rangle$ al autoket normalizado de todas las n pertenecientes a los autovalores cero. Es decir,

$$n_a|0\rangle = 0,$$

y, por lo tanto, de (12′)

$$\langle 0|\eta_a \bar{\eta}_a|0\rangle = 0.$$

Luego,

$$\bar{\eta}_a|0\rangle = 0, \tag{15′}$$

correspondiente a (15). Además

$$\langle 0|\bar{\eta}_a \eta_a|0\rangle = \langle 0|(1 - n_a)|0\rangle = \langle 0|0\rangle = 1,$$

y por tanto, $\eta_a|0\rangle$ está normalizado, y

$$n_a \eta_a|0\rangle = \eta_a \bar{\eta}_a \eta_a|0\rangle = \eta_a (1 - n_a)|0\rangle = \eta_a|0\rangle,$$

o sea, $\eta_a|0\rangle$ es un autoket de n_a perteneciente al autovalor uno. Además, es un autoket de las otras n perteneciente a los autovalores cero, puesto que las otras n conmutan con η_a. Generalizando el argumento vemos que $\eta_a \eta_b \eta_c \ldots \eta_g|0\rangle$ está normalizado y es un autoket común a todas las n que pertenece a los autovalores uno de n_a, n_b, n_c, ... n_g y cero de las otras n. Esto nos permite poner

$$A|\alpha^a \alpha^b \alpha^c \ldots \alpha^g\rangle = \eta_a \eta_b \eta_c \ldots \eta_g|0\rangle, \tag{17′}$$

en donde ambos miembros son antisimétricos con respecto a los símbolos a, b, c, ..., g. Corresponde a (17).

Si consideramos un conjunto diferente de kets básicos $|\beta^A\rangle$ para un fermión, podemos introducir un nuevo conjunto de operadores lineales η_A en correspondencia con ellos. Encontramos entonces, razonando igual que hicimos con los bosones, que las nuevas η_A están relacionadas con las originales por (21). Vemos con ello que hay un procedimiento de segunda cuantificación para fermiones, similar al de bosones, con la única diferen-

cia de que para fermiones han de emplearse las relaciones de conmutación (11′) que reemplazan a las relaciones de conmutación (11) que se aplican para bosones.

Un operador lineal simétrico U_T de la forma (22) puede expresarse en función de las variables η, $\bar{\eta}$ por un método análogo al empleado para bosones. La ecuación (24) continúa siendo válida e igualmente la (25) reemplazando S por A. En vez de (26) tenemos ahora

$$U_T \bar{\eta}_{x_1} \eta_{x_2} \ldots |0\rangle = \sum_a \sum_r (-)^{r-1} \eta_a \eta_{x_r}^{-1} \eta_{x_1} \eta_{x_2} \ldots |0\rangle \langle a|U|x_r\rangle$$

$$= \sum_{ab} \eta_a \sum_r (-)^{r-1} \eta_{x_r}^{-1} \eta_{x_1} \eta_{x_2} \ldots |0\rangle \, \delta_{bx_r} \langle a|U|b\rangle, \qquad (26')$$

$\eta_{x_r}^{-1}$ significa que el factor η_{x_r} ha de eliminarse sin cambiar su posición entre las otras η_x antes de la eliminación. En vez de (27) tenemos

$$\bar{\eta}_b \bar{\eta}_{x_1} \bar{\eta}_{x_2} \ldots |0\rangle = \sum_r (-)^{r-1} \eta_{x_r}^{-1} \eta_{x_1} \eta_{x_2} \ldots |0\rangle \delta_{bx_r}, \qquad (27')$$

luego, (28) permanece inalterada y, como consecuencia, (29) también. Para U_T en el caso de fermiones tenemos la misma forma final (29) que en el caso de bosones. De manera análoga, un operador lineal simétrico V_T de la forma (30) puede expresarse como

$$V_T = \sum_{abcd} \eta_a \eta_b \langle ab|V|cd\rangle \bar{\eta}_d \bar{\eta}_c, \qquad (35')$$

que es una de las formas de escribir (35).

El desarrollo precedente nos muestra que hay una profunda analogía entre la teoría de fermiones y la de bosones, necesitándose solamente ligeros cambios en las ecuaciones generales del formalismo al pasar de una a otra.

Sin embargo, hay un desarrollo de la teoría de fermiones que no tiene su análogo para bosones. Para fermiones solamente hay las dos alternativas de ocupación o no ocupación de un estado y *hay simetría entre ambas alternativas*. Puede demostrarse la simetría matemáticamente haciendo una transformación que intercambie los conceptos de "ocupado" y "desocupado", a saber

$$\eta_a^* = \bar{\eta}_a, \qquad \bar{\eta}_a^* = \eta_a,$$

$$n_a^* = \eta_a^* \bar{\eta}_a^* = 1 - n_a.$$

Los operadores de creación de las variables sin estrella son los operadores de aniquilación de las variables con estrella y viceversa. Vemos que las variables con estrella satisfacen ahora las mismas condiciones cuánticas y tienen exactamente las mismas propiedades que las variables sin estrella.

Si sólo hay unos pocos estados desocupados, un ket standard conveniente para trabajar sería aquel para el cual todos los estados están ocupados, es decir, el ket $|0^*\rangle$ que verifica

$$n_a|0^*\rangle = |0^*\rangle.$$

y que, por tanto, verifica también

$$n_a^* \, |0^*\rangle = 0,$$

o sea,

$$\bar{\eta}_a^* \, |0^*\rangle = 0.$$

Los otros estados para el conjunto vendrán representados ahora por

$$\bar{\eta}_a^* \, \bar{\eta}_b^* \, \eta_c^* \ldots |0^*\rangle,$$

en la cual las variables que aparecen se refieren a los estados fermiónicos sin ocupar a, b, c ... Podemos considerar estos estados fermiónicos desocupados como huecos entre los estados ocupados y las variables $\bar{\eta}_g^*$ como los operadores de creación de tales huecos. Los huecos tienen la misma existencia física que las partículas originales y son también fermiones.

XI

TEORÍA RELATIVISTA DEL ELECTRÓN

66. *Teoría relativista de una partícula*

La teoría que hemos ido construyendo, hasta ahora, es esencialmente una teoría no relativista. Hemos trabajado siempre en un sistema de referencia de Lorentz particular y hemos construido la teoría en analogía con la dinámica clásica no relativista. Tratemos ahora de hacer invariante la teoría frente a las transformaciones de Lorentz, de modo que cumpla con el principio de la relatividad restringida. Ello es necesario para poder aplicar la teoría a partículas de gran velocidad. No necesitamos que la teoría satisfaga la relatividad general, puesto que la relatividad general sólo es necesaria cuando se trabaja con la gravitación, y las fuerzas gravitatorias son totalmente despreciables en los fenómenos atómicos.

Veamos cómo pueden adaptarse las ideas básicas de la teoría cuántica al principio relativista según el cual las cuatro dimensiones del espacio-tiempo deben tratarse de igual forma. El principio general de superposición de estados tal como se enunció en el capítulo I es un principio relativista, puesto que se aplica a 'estados' con el significado relativista del espacio-tiempo. Sin embargo, el concepto general de observable no encaja bien, puesto que un observable puede contener entes físicos en puntos muy separados en un mismo instante de tiempo. Por consiguiente, si se trabaja con una representación general referida a un conjunto completo de observables que conmutan, la teoría no puede mostrar la simetría entre espacio y tiempo exigida por la relatividad. En la mecánica cuántica relativista debemos contentarnos con una sola representación que goce de esta simetría. Existe, sin embargo, la posibilidad de transformarla a otra representación referida a otro sistema de referencia de Lorentz particular, si ello resulta útil.

En el problema de una sola partícula, para expresar la simetría entre espacio y tiempo, hemos de usar la representación de Schrödinger. Pongamos x_1, x_2, x_3 en lugar de x, y, z y x_0 en lugar de ct. La función de onda dependiente del tiempo aparece entonces en la forma $\psi(x_0\, x_1\, x_2\, x_3)$ y nos proporciona una base para tratar las cuatro x de la misma manera.

Emplearemos la notación relativista escribiendo las cuatro x como x_μ ($\mu = 0, 1, 2, 3$). Cualquier vector del espacio-tiempo que se transforme en una transformación de Lorentz como los cuatro elementos dx_μ se escribirá a_μ con un subíndice griego. Podemos subir el índice de acuerdo con las reglas

$$a^0 = a_0, \quad a^1 = -a_1, \quad a^2 = -a_2, \quad a^3 = -a_3. \tag{1}$$

Las a_μ se llaman componentes contravariantes del vector a, y las a^μ son las componentes covariantes. Dos vectores a_μ y b_μ tienen un producto escalar invariante bajo las transformaciones de Lorentz igual a

$$a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3 = a^\mu b_\mu = a_\mu b^\mu.$$

Debe entenderse que un índice no numérico repetido indica sumación. El tensor fundamental $g^{\mu\nu}$ está definido por

$$\left.\begin{aligned} g^{00} = 1, \quad g^{11} = g^{22} = g^{33} = -1, \\ g^{\mu\nu} = 0 \quad \text{para } \mu \neq \nu. \end{aligned}\right\} \tag{2}$$

Con su ayuda, las reglas (1) que relacionan los componentes covariantes y contravariantes se pueden escribir

$$a^\mu = g^{\mu\nu} a_\nu.$$

En la representación de Schrödinger el momento, cuyas componentes escribiremos p_1, p_2, p_3 en vez de p_x, p_y, p_z, es igual al operador

$$p_r = -i\hbar\, \partial/\partial x_r \qquad (r = 1, 2, 3). \tag{3}$$

Ahora bien, los cuatro operadores $\partial/\partial x_\mu$ forman las componentes covariantes de un cuadrivector cuyas componentes contravariantes se escriben $\partial/\partial x^\mu$. Para introducir (3) en la teoría relativista hemos de escribirla primeramente con los índices cambiados

$$p_r = i\hbar\, \partial/\partial x^r,$$

y extenderla entonces a la ecuación completa del cuadrivector

$$p_\mu = i\hbar\, \partial/\partial x^\mu. \tag{4}$$

Tenemos que introducir, por lo tanto, una nueva variable dinámica p_0 igual al operador $i\hbar\, \partial/\partial x_0$. Puesto que combinada con los momentos p_r forma un cuadrivector, tiene que tener el significado físico de la energía de la partícula dividida por c. Podemos seguir el desarrollo de la teoría considerando las cuatro p en pie de igualdad con las cuatro x.

En la teoría del electrón que vamos a desarrollar tendremos que introducir un nuevo grado de libertad que da cuenta de un movimiento interno del electrón. La función de onda tendrá que incluir, pues, una nueva variable además de las cuatro x.

67. *La ecuación de onda para el electrón*

Consideremos en primer lugar el caso del movimiento de un electrón en ausencia de campo electromagnético, es decir, el problema simple de una partícula libre ya estudiado en § 30 con la posible adición de grados de libertad internos. El hamiltoniano relativista en mecánica clásica viene dado por la ecuación (23) de § 30 y conduce a la ecuación de ondas

$$\{p_0 - (m^2c^2 + p_1^2 + p_2^2 + p_3^2)^{\frac{1}{2}}\}\psi = 0, \tag{5}$$

donde las p se interpretan como operadores de acuerdo con (4). La ecuación (5) si bien tiene en cuenta la relación entre energía y momento que exige la relatividad, no es del todo satisfactoria desde el punto de vista de la teoría relativista, por la asimetría entre p_0 y las otras p, y debido a esto no puede generalizarse de una manera relativista para el caso de que exista un campo. Hemos de buscar, por consiguiente, una nueva ecuación de ondas.

Si multiplicamos a la izquierda la ecuación de ondas (5) por el operador $\{p_0 + (m^2c^2 + p_1^2 + p_2^2 + p_3^2)^{\frac{1}{2}}\}$ obtenemos la ecuación

$$\{p_0^2 - m^2c^2 - p_1^2 - p_2^2 - p_3^2\}\psi = 0, \tag{6}$$

que es de una forma invariante relativista y, por consiguiente, es más adecuada como base de una teoría relativista. La ecuación (6) no es completamente equivalente a la ecuación (5) pues si bien toda solución de (5) es también solución de (6) el recíproco no es cierto. Sólo las soluciones de (6) que pertenecen a valores positivos de p_0 son también soluciones de (5).

La ecuación de ondas (6) no es de la forma requerida por las leyes generales de la teoría cuántica, puesto que es cuadrática en p_0. En § 27 dedujimos mediante argumentos totalmente generales que la ecuación de ondas ha de ser lineal en el operador $\partial/\partial t$ o p_0, como se tiene en la ecuación (7) de esta sección. Buscamos, por lo tanto, una ecuación de ondas que sea lineal en p_0 y aproximadamente equivalente a la (6). A fin de que esta ecuación se transforme de una manera sencilla bajo una transformación de Lorentz, intentaremos modificarla para que sea racional y lineal no sólo en p_0, sino también en p_1, p_2 y p_3, y, por lo tanto, de la forma

$$\{p_0 - \alpha_1 p_1 - \alpha_2 p_2 - \alpha_3 p_3 - \beta\}\psi = 0, \tag{7}$$

donde las α y β son independientes de las p. Puesto que estamos considerando el caso de que no exista ningún campo, todos los puntos del espacio-tiempo han de ser equivalentes; luego, el operador de la ecuación de ondas no ha de contener a las x. Por consiguiente, las α y β han de independientes también de las x y, por lo tanto, han de conmutar con las p y las x. En consecuencia, describen algún nuevo grado de libertad relacionado

con algún movimiento interno del electrón. Veremos más adelante que dan cuenta del spin del electrón.

Multiplicando (7) a la izquierda por el operador $\{p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta\}$ tenemos

$$\left\{ p_0^2 - \sum_{123} [\alpha_1^2 p_1^2 + (\alpha_1 \alpha_2 + \alpha_2 \alpha_1) p_1 p_2 + (\alpha_1 \beta + \beta\alpha_1) p_1] - \beta^2 \right\} \psi = 0,$$

donde $\sum_{123}$ se refiere a permutaciones circulares de los subíndices 1, 2, 3. Esta ecuación coincide con (6) si las α y β satisfacen las relaciones

$$\begin{aligned} \alpha_1^2 &= 1, & \alpha_1 \alpha_2 + \alpha_2 \alpha_1 &= 0, \\ \beta^2 &= m^2c^2, & \alpha_1 \beta + \beta\alpha_1 &= 0, \end{aligned}$$

y las relaciones que se obtienen de éstas al permutar los índices 1, 2, 3. Escribiendo

$$\beta = \alpha_m mc,$$

podemos resumir estas relaciones en la ecuación única

$$\alpha_a \alpha_b + \alpha_b \alpha_a = 2\delta_{ab} \quad (a, b = 1, 2, 3, \text{ o bien } m). \tag{8}$$

Las cuatro α anticonmutan todas entre sí y el cuadrado de cada una de ellas es la unidad.

Por lo tanto, atribuyendo propiedades adecuadas a las α y β, podemos hacer la ecuación de ondas (7) equivalente a la (6) en lo que concierne al movimiento del electrón como un todo. Supongamos ahora que (7) es la ecuación de ondas relativista para el movimiento de un electrón en ausencia de campo. Esto, sin embargo, origina una dificultad debida al hecho que (7), como (6), no es exactamente equivalente a (5), y permite soluciones correspondientes tanto a valores negativos como a valores positivos de p_0. Las primeras no corresponden, por supuesto, a ningún movimiento realmente observable del electrón. Por ahora consideraremos solamente las soluciones de energía positiva y dejaremos la discusión de las de energía negativa para § 73.

Podemos obtener fácilmente una representación de las cuatro α. Tienen propiedades algebraicas similares a las σ introducidas en § 37, que podían representarse por matrices con dos filas y dos columnas. Con matrices de dos filas y dos columnas no podemos obtener una representación de más de tres cantidades que anticonmuten; necesitamos cuatro filas y cuatro columnas para obtener una representación de las cuatro α que anticonmutan. Conviene expresar previamente las α en función de las σ y de un segundo conjunto de tres variables que anticonmutan ρ_1, ρ_2, ρ_3 y que son independientes de las σ y que conmutan, por tanto, con ellas y cuyos cuadrados valen la unidad. Podemos tomar entre otras posibilidades

$$\alpha_1 = \rho_1 \sigma_1, \qquad \alpha_2 = \rho_1 \sigma_2, \qquad \alpha_3 = \rho_1 \sigma_3, \qquad \alpha_m = \rho_3, \tag{9}$$

Las α satisfarán entonces todas las relaciones (8), como puede comprobarse fácilmente. Si tomamos ahora una representación con ρ_3 y σ_3 diagonales, obtendremos el siguiente esquema de matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\rho_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad \rho_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \quad \rho_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Observemos que las ρ y las σ son hermíticas, lo que hace que las α también lo sean.

En correspondencia con las cuatro filas y las cuatro columnas, la función de onda ψ ha de contener una variable que tome cuatro valores a fin de que las matrices puedan multiplicarse por ella. Otra posibilidad es considerar que la función de ondas tiene cuatro componentes, cada una función solamente de las cuatro x. Vimos en § 37 que el spin del electrón exige que la función de onda tenga dos componentes. El hecho de que nuestra presente teoría dé lugar a cuatro, es debido a que la función de ondas tiene doble número de soluciones de las que debería tener, correspondiendo la mitad de ellas a estados de energía nagtiva.

Con la ayuda de (9) la ecuación de onda (7) puede escribirse en notación vectorial de tres dimensiones

$$\{p_0 - \rho_1 (\boldsymbol{\sigma}, \mathbf{p}) - \rho_3 mc\}\psi = 0. \tag{10}$$

Para generalizar esta ecuación al caso de que haya un campo electromagnético, seguimos la regla clásica de substituir p_0 y $\mathbf{p}$ por $p_0 + e/c \cdot A_0$ y $\mathbf{p} + e/c \cdot \mathbf{A}$, donde A_0 y $\mathbf{A}$ son los potenciales escalar y vector del campo en el lugar donde se halla el electrón. Así obtenemos la ecuación

$$\left\{ p_0 + \frac{e}{c} A_0 - \rho_1 \left(\boldsymbol{\sigma}, \mathbf{p} + \frac{e}{c} \mathbf{A} \right) - \rho_3 mc \right\} \psi = 0, \tag{11}$$

que es la ecuación de onda fundamental de la teoría relativista del electrón.†

Las cuatro componentes de ψ en (10) u (11) han de representarse escritas una debajo de otra, formando una matriz de una sola columna. Las matrices cuadradas ρ y σ quedan multiplicadas entonces por la matriz de una sola columna ψ según la regla ordinaria de multiplicación de matri-

† *N. de T.:* Esta ecuación se conoce con el nombre de *ecuación de Dirac* para el electrón, en honor a su descubridor.

ces, siendo el producto otra matriz de una sola columna. La función de onda imaginaria conjugada que representa un bra debe representarse con sus cuatro componentes una al lado de otra formando una matriz de una sola fila, que puede multiplicarse a la derecha por una matriz cuadrada para dar otra matriz de una sola fila. A esta función de onda imaginaria conjugada, representada por una matriz de una sola fila la designaremos $\bar{\psi}^\dagger$, donde el signo $\dagger$ indica la traspuesta de una matriz, es decir, el resultado de intercambiar filas por columnas. La imaginaria conjugada de la ecuación (11) será entonces

$$\bar{\psi}^\dagger \left\{ p_0 + \frac{e}{c} A_0 - \rho_1 \left(\boldsymbol{\sigma}, \mathbf{p} + \frac{e}{c} \mathbf{A} \right) - \rho_3 mc \right\} = 0, \tag{12}$$

con los operadores p actuando a la izquierda. Los operadores de derivación actuando a la izquierda han de interpretarse de acuerdo con (24) de § 22.

68. *Invariancia bajo una transformación de Lorentz*

Antes de pasar a discutir las consecuencias físicas de la ecuación de onda (11) o (12) vamos a comprobar previamente que nuestra teoría es invariante bajo las transformaciones de Lorentz, o más exactamente, que los resultados físicos a que conduce la teoría son independientes del sistema de referencia de Lorentz considerado. Esto no se desprende inmediatamente de la forma de la ecuación de onda (11). Hemos de comprobar que, al escribir la ecuación de onda en un sistema de Lorentz distinto, las soluciones de la nueva ecuación de onda pueden ponerse en correspondencia biyectiva con las de la primitiva de tal modo que las soluciones correspondientes representen el mismo estado. Para cada sistema de Lorentz, el cuadrado del módulo de la función de onda, que se obtiene sumando los de las cuatro componentes, tiene que ser igual a la probabilidad por unidad de volumen de que el electrón esté en un cierto lugar de esta representación. Podemos denominarla *densidad de probabilidad.* Sus valores, calculados en diferentes sistemas de Lorentz para funciones de onda que representen el mismo estado, han de estar relacionados como las componentes temporales de un cuadrivector en estos sistemas. Además la divergencia cuadridimensional de este cuadrivector ha de anularse, significando con ello conservación del electrón, es decir, que el electrón no puede parecer o desaparecer en un volumen sin pasar por su contorno.

Para abreviar es conveniente introducir el símbolo $\alpha_0 = 1$ y suponer que los subíndices de las cuatro α_μ ($\mu = 0, 1, 2, 3$) pueden subirse de acuerdo con las reglas (1), a pesar de que estas cuatro α no formen las

componentes de ningún cuadrivector. Podemos escribir ahora la ecuación de onda (11)

$$\{\alpha^{\mu}(p_{\mu} + e/c \cdot A_{\mu}) - \alpha_m mc\}\psi = 0. \tag{13}$$

Las cuatro α^{μ} verifican

$$\alpha^{\mu}\alpha_m \alpha^{\nu} + \alpha^{\nu}\alpha_m \alpha^{\mu} = 2g^{\mu\nu}\alpha_m \tag{14}$$

con las $g^{\mu\nu}$ definidas en (2), como puede comprobarse tomando separadamente los casos en que μ y ν son ambas 0, cuando una de ellas es 0, y cuando ninguna lo es.

Apliquemos una transformación de Lorentz infinitesimal y distigamos mediante una estrella las cantidades referidas al nuevo sistema de referencia. Las componentes del cuadrivector p_{μ} se transformarán de acuerdo con ecuaciones del tipo

$$p^*_{\mu} = p_{\mu} + a_{\mu}{}^{\nu}p_{\nu}, \tag{15}$$

donde las $a_{\mu}{}^{\nu}$ son infinitésimos de primer orden. Despreciaremos las cantidades cuadráticas en las a ya que son de segundo orden. La condición para que una transformación sea de Lorentz es que

$$p^*_{\mu}p^{\mu *} = p_{\mu}p^{\mu},$$

que da

$$a_{\mu}{}^{\nu}p_{\nu}p^{\mu} + p_{\mu}a^{\mu\nu}p_{\nu} = 0,$$

y por lo tanto

$$a^{\mu\nu} + a^{\nu\mu} = 0. \tag{16}$$

Las componentes de A_{μ} se transformarán de acuerdo con la misma ley, y así tenemos

$$p_{\mu} + e/c \cdot A_{\mu} = p^*_{\mu} + e/c \cdot A^*_{\mu} - a_{\mu}{}^{\nu}(p^*_{\nu} + e/c \cdot A^*_{\nu}).$$

La ecuación de onda (13) se transforma pues en

$$\{(\alpha^{\mu} - \alpha^{\lambda}a_{\lambda}{}^{\mu})(p^*_{\mu} + e/c \cdot A^*_{\mu}) - a_m mc\}\psi = 0. \tag{17}$$

Definimos ahora

$$M = \tfrac{1}{4}a_{\rho\sigma}\alpha^{\rho}\alpha_m\alpha^{\sigma}. \tag{18}$$

Luego, según (14)

$$\begin{aligned}\alpha^{\mu}\alpha_m M - M\,\alpha_m\alpha^{\mu} &= \tfrac{1}{4}a_{\rho\sigma}\,\{(\alpha^{\mu}\alpha_m\alpha^{\rho} + \alpha^{\rho}\alpha_m\alpha^{\mu})\alpha_m\alpha^{\sigma} - \\ &\qquad - \alpha^{\rho}\alpha_m(\alpha^{\mu}\alpha_m\alpha^{\sigma} + \alpha^{\sigma}\alpha_m\alpha^{\mu})\} \\ &= \tfrac{1}{2}a_{\rho\sigma}(g^{\mu\rho}\alpha^{\sigma} - \alpha^{\rho}g^{\mu\sigma}) \\ &= -a_{\rho}{}^{\mu}\alpha^{\rho}\end{aligned}$$

con la ayuda de (16). De aquí

$$\alpha^\mu(1 + \alpha_m M) = (1 + M\alpha_m)(\alpha^\mu - a_\rho{}^\mu \alpha^\rho). \tag{19}$$

Multiplicando (17) por $(1 + M\alpha_m)$ a la izquierda, obtenemos

$$\{\alpha^\mu(1 + \alpha_m M)(p^*_\mu + e/c \cdot A^*_\mu) - (\alpha_m + M)mc\}\psi = 0.$$

Y si ponemos

$$(1 + \alpha_m M)\psi = \psi^*, \tag{20}$$

resulta

$$\{\alpha^\mu(p^*_\mu + e/c \cdot A^*_\mu) - \alpha_m mc\}\psi^* = 0. \tag{21}$$

que es de la misma forma que (13) en las variables con estrella p^*_μ, A^*_μ, ψ^*, y muestra que (13) es invariante bajo una transformación de Lorentz infinitesimal, siempre y cuando ψ esté sujeta a la transformación correcta dada por (20). Mediante transformaciones de Lorentz infinitesimales podemos construir una transformación de Lorentz finita, luego, la ecuación de onda (13) también es invariante bajo una transformación de Lorentz finita. Obsérvese que las matrices α^μ quedan totalmente inalteradas.

La invariancia que hemos comprobado significa que las soluciones ψ de la primitiva ecuación de onda (13) están en correspondencia biyectiva con las soluciones ψ^* de la nueva ecuación de onda (21), estando relacionadas las soluciones correspondientes por (20). Vamos a suponer que las soluciones correspondientes representan el mismo estado físico. Hemos de comprobar ahora que las interpretaciones físicas de las soluciones correspondientes, referidas a sus respectivos sistemas de referencia de Lorentz, concuerdan. Esto exige que $\bar{\psi}^\dagger\psi$ dé la densidad de probabilidad referida al sistema primitivo y $\bar{\psi}^{*\dagger}\psi^*$ la densidad de probabilidad referida al nuevo sistema. Examinemos la relación entre estas dos cantidades. $\bar{\psi}^\dagger\psi$ es lo mismo que $\bar{\psi}^\dagger\alpha^0\psi$ que forma una de las cuatro cantidades $\bar{\psi}^\dagger\alpha^\mu\psi$ que consideraremos conjuntamente.

Las ecuaciones (18) y (16) muestran que M es imaginario puro. Luego, la conjugada imaginaria de la ecuación (20) es

$$\bar{\psi}^{*\dagger} = \bar{\psi}^\dagger(1 - M\,\alpha_m).$$

De donde

$$\begin{aligned}\bar{\psi}^{*\dagger}\alpha^\mu\psi^* &= \bar{\psi}^\dagger(1 - M\,\alpha_m)\alpha^\mu(1 + \alpha_m M)\psi \\ &= \bar{\psi}^\dagger(1 - M\,\alpha_m)(1 + M\,\alpha_m)(\alpha^\mu - a_\nu{}^\mu\alpha^\nu)\psi\end{aligned}$$

según (19). Esto se reduce con la ayuda de (16), a

$$\begin{aligned}\bar{\psi}^{*\dagger}\alpha^\mu\psi^* &= \bar{\psi}^\dagger(\alpha^\mu - a_\nu{}^\mu\alpha^\nu)\psi \\ &= \bar{\psi}^\dagger\alpha^\mu\psi + a^\mu{}_\nu\bar{\psi}^\dagger\alpha^\nu\psi\end{aligned}$$

$\{\alpha^\mu(p_\mu + \frac{e}{c} A_\mu) - \alpha_m mc\} \psi = 0$ (13)

Si bajamos el índice μ obtenemos una ecuación de la misma forma que (15), lo que muestra que las cuatro cantidades $\psi^\dagger \alpha_\mu \psi$ se transforman como las componentes contravariantes de un cuadrivector. Luego, $\psi^\dagger \psi$ se transforma como la componente temporal de un cuadrivector, que es la ley de transformación correcta de la densidad de probabilidad. Si multiplicamos por c las componentes espaciales del cuadrivector, o sea, las $\psi^\dagger \alpha_r \psi$, obtenemos la corriente de probabilidad o probabilidad de que el electrón atraviese la unidad de área por unidad de tiempo.

Observemos que $\psi^\dagger \alpha_m \psi$ es invariante, puesto que

$$\begin{aligned} \psi^{*\dagger} \alpha_m \psi^* &= \psi^\dagger (1 - M\,\alpha_m) \alpha_m (1 + \alpha_m M) \psi \\ &= \psi^\dagger \alpha_m \psi. \end{aligned}$$

Hemos de verificar finalmente la ley de conservación, es decir, que la divergencia

$$\frac{\partial}{\partial x_\mu} (\psi^\dagger \alpha_\mu \psi) \tag{22}$$

es nula. Para probarlo multiplicaremos la ecuación (13) por $\psi^\dagger$ a la izquierda. El resultado es

$$\psi^\dagger \alpha^\mu \left(i\hbar \frac{\partial \psi}{\partial x^\mu} + \frac{e}{c} A_\mu \psi \right) - \psi^\dagger \alpha_m mc \psi = 0.$$

La ecuación imaginaria conjugada es

$$\left(-i\hbar \frac{\partial \psi^\dagger}{\partial x^\mu} + \psi^\dagger \frac{e}{c} A_\mu \right) \alpha^\mu \psi - \psi^\dagger \alpha_m mc \psi = 0.$$

Restando y dividiendo por $i\hbar$, obtenemos

$$\psi^\dagger \alpha^\mu \frac{\partial \psi}{\partial x^\mu} + \frac{\partial \psi^\dagger}{\partial x^\mu} \alpha^\mu \psi = 0,$$

que expresa precisamente la anulación de (22). Completamos de esta manera la prueba de que nuestra teoría da resultados concordantes en cualquier sistema de referencia a que se aplique.

69. *Movimiento de un electrón libre*

Es interesante considerar el movimiento de un electrón libre en la imagen de Heisenberg de acuerdo con la teoría anterior y estudiar las ecuaciones del movimiento de Heisenberg. Estas ecuaciones del movimiento pueden integrarse exactamente, como hizo por primera vez Schrö-

(*) Recordar que las ec. de Heisenberg son:

$$\frac{dN_t}{dt} = [N_t, \hat{H}_t] \rightarrow$$

dinger.‡ Omitiremos para abreviar el subíndice t que según la notación de § 28 tendríamos que añadir a las variables dinámicas que varían con el tiempo en la imagen de Heisenberg.

Hemos de tomar como hamiltoniano la expresión que se obtiene para cp_0 al igualar a cero el operador que actúa sobre ψ en (10), es decir,

$$H = c\rho_1(\boldsymbol{\sigma}, \mathbf{p}) + \rho_3 mc^2 = c(\boldsymbol{\alpha}, \mathbf{p}) + \rho_3 mc^2. \tag{23}$$

$[H, p]$

Vemos inmediatamente que el momento conmuta con H y que, por lo tanto, es una constante del movimiento. Además la componente x_1 de la velocidad es

$$\dot{x}_1 = [x_1, H] = c\alpha_1. \tag{24}$$

Este resultado es bastante sorprendente, pues es una relación totalmente diferente entre velocidad y momento de la que se tiene en mecánica clásica. Sin embargo, está relacionada con la expresión $\psi^\dagger c\alpha_1\psi$ de una componente de la corriente de probabilidad. La $\dot{x}_1$ dada por (24) tiene los autovalores $\pm c$ correspondientes a los autovalores ± 1 de α_1. Como $\dot{x}_2$ y $\dot{x}_3$ son análogas, podemos concluir que *la medición de una componente de la velocidad de un electrón da con certeza el resultado* $\pm c$. Puede verse fácilmente que esta conclusión vale también cuando existe campo.

Como en la práctica se observa que los electrones tienen velocidades considerablemente menores que la de la luz, podría parecer que tenemos una contradicción con la experiencia. Sin embargo, la contradicción no es real puesto que la velocidad teórica de la conclusión anterior es la velocidad en un instante, mientras que las velocidades observadas son siempre promedios de velocidad durante intervalos de tiempo apreciables. Examinando con más detalle las ecuaciones del movimiento encontraremos que la velocidad no es en absoluto constante, sino que oscila rápidamente alrededor de un valor medio que concuerda con el valor observado.

Puede justificarse fácilmente que una medida de una componente de la velocidad ha de conducir en una teoría relativista al resultado $\pm c$, considerando simplemente el principio de incertidumbre de § 24. Para medir la velocidad hemos de medir la posición en dos instantes de tiempo ligeramente diferentes y dividir entonces el cambio de posición por el intervalo de tiempo. (No tendría valor alguno medir el momento y aplicar una fórmula, puesto que la relación ordinaria entre velocidad y momento no es válida.) Para que la velocidad que medimos pueda aproximarse a la velocidad instantánea, el intervalo de tiempo entre las dos medidas de la posición ha de ser muy corto y, por lo tanto, las medidas muy precisas. La gran precisión con que debe conocerse la posición del electrón durante el intervalo de tiempo, ha de dar lugar, según el princi-

‡ SCHRÖDINGER, *Sitzungsb. d. Berlin. Akad.*, 1930, p. 418.

→ el símbolo t como subíndice [illegible] pero resulta el hecho que los op. dependen explícitamente del tiempo.

pio de incertidumbre, a una indeterminación casi completa en su momento. Esto significa que casi todos los valores del momento son igualmente probables, luego, el momento es con casi toda probabilidad infinito. Un valor infinito de una componente del momento corresponde al valor $\pm c$ de la componente de la velocidad correspondiente.

Examinemos ahora cómo varía la velocidad del electrón con el tiempo. Tenemos

$$i\hbar\dot{\alpha}_1 = \alpha_1 H - H\alpha_1.$$

Como α_1 anticonmuta con todos los términos de H excepto $c\alpha_1 p_1$, se tendrá

$$\alpha_1 H + H\alpha_1 = \alpha_1 c\alpha_1 p_1 + c\alpha_1 p_1 \alpha_1 = 2cp_1,$$

de donde

$$\left.\begin{aligned} i\hbar\dot{\alpha}_1 &= 2\alpha_1 H - 2cp_1, \\ &= -2H\alpha_1 + 2cp_1. \end{aligned}\right\} \quad (25)$$

Como H y p_1 son constantes, de la primera ecuación (25) se deduce que

$$i\hbar\ddot{\alpha}_1 = 2\dot{\alpha}_1 H. \quad (26)$$

Esta ecuación diferencial en $\dot{\alpha}_1$ puede integrarse inmediatamente obteniéndose

$$\dot{\alpha}_1 = \dot{\alpha}_1^0 e^{-2iHt/\hbar}, \quad (27)$$

donde $\dot{\alpha}_1^0$ es una constante, igual al valor de $\dot{\alpha}_1$ para $t = 0$. El factor $e^{-2iHt/\hbar}$ tiene que colocarse a la derecha del factor $\dot{\alpha}_1^0$ en (27) porque la H aparece a la derecha de $\dot{\alpha}_1{}^0$ en (26). La segunda ecuación (25) conduce del mismo modo al resultado

$$\dot{\alpha}_1 = e^{2iHt/\hbar}\dot{\alpha}_1{}^0.$$

Podemos completar ahora fácilmente la integración de la ecuación de movimiento en x_1. De (27) y de la primera ecuación (25)

$$\alpha_1 = \tfrac{1}{2}i\hbar\dot{\alpha}_1^0 e^{-2iHt/\hbar}H^{-1} + cp_1 H^{-1}, \quad (28)$$

de donde la integral respecto al tiempo de la ecuación (24) es

$$x_1 = -\tfrac{1}{4}c\hbar^2\dot{\alpha}_1^0 e^{-2iHt/\hbar}H^{-2} + c^2 p_1 H^{-1}t + a_1, \quad (29)$$

siendo α_1 una constante.

Vemos en (28) que la componente x_1 de la velocidad, o sea $c\alpha_1$, consta de dos partes; una parte constante $c^2 p_1 H^{-1}$, relacionada con el momento por la fórmula clásica relativista, y una parte oscilatoria

$$\tfrac{1}{2}ic\hbar\dot{\alpha}_1^0 e^{-2iHt/\hbar}H^{-1},$$

de alta frecuencia $2H/h$, que vale como mínimo $2mc^2/h$. En una medición práctica de la velocidad sólo se observaría la parte constante, pues

la medición sólo puede determinar la velocidad promedio durante un intervalo de tiempo mucho menor que $h/2mc^2$. La parte oscilatoria hace que el valor instantáneo de $\dot{x}_1$ tenga los autovalores $\pm c$. La parte oscilatoria de x_1 es pequeña; según (29) vale

$$-\tfrac{1}{4}c\hbar^2\dot{\alpha}_1^0 e^{-2iHt/\hbar}H^{-2} = \tfrac{1}{2}ic\hbar(\alpha_1 - cp_1H^{-1})H^{-1},$$

que es del orden de magnitud de $\hbar/mc$, ya que $(\alpha_1 - cp_1H^{-1})$ es del orden de magnitud unidad.

70. *Existencia del spin*

Vimos en § 67 que la ecuación de onda correcta para el electrón en ausencia de un campo electromagnético, es decir, la ecuación (7) o (10) es equivalente a la ecuación de onda (6) sugerida por analogía con la teoría clásica. Esta equivalencia deja de ser válida cuando hay un campo. La ecuación de onda que puede esperarse en este caso por analogía con la teoría clásica es

$$\left\{\left(p_0 + \frac{e}{c}A_0\right)^2 - \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - m^2c^2\right\}\psi = 0, \tag{30}$$

en la cual el operador es precisamente el hamiltoniano relativista clásico. Si, para obtener a partir de (11) una expresión que se parezca a (30) tanto como sea posible, la multiplicamos a la izquierda por el factor

$$p_0 + \frac{e}{c}A_0 + \rho_1\left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right) + \rho_3\, mc,$$

obtenemos

$$\left\{\left(p_0 + \frac{e}{c}A_0\right)^2 - \left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - m^2c^2 - \rho_1 \right.$$
$$\left.\left[\left(p_0 + \frac{e}{c}A_0\right)\left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right) - \left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right)\left(p_0 + \frac{e}{c}A_0\right)\right]\right\}\psi = 0. \tag{31}$$

Empleemos ahora la fórmula general, válida para dos vectores tridimensionales $\mathbf{B}$ y $\mathbf{C}$ cualesquiera que conmuten con $\boldsymbol{\sigma}$,

$$(\boldsymbol{\sigma},\ \mathbf{B})(\boldsymbol{\sigma},\ \mathbf{C}) = \sum_{123} \{\sigma_1{}^2\, B_1\, C_1 + \sigma_1\, \sigma_2\, B_1\, C_2 + \sigma_2\, \sigma_1\, B_2\, C_1\},$$

donde la sumación está extendida a las permutaciones circulares de los subíndices 1, 2, 3, y que también puede escribirse

$$\begin{aligned}(\boldsymbol{\sigma},\ \mathbf{B})(\boldsymbol{\sigma},\ \mathbf{C}) &= (\mathbf{B},\ \mathbf{C}) + i \sum_{123} \sigma_3(B_1\, C_2 - B_2\, C_1)\\ &= (\mathbf{B},\ \mathbf{C}) + i(\boldsymbol{\sigma},\ \mathbf{B} \times \mathbf{C}).\end{aligned} \tag{32}$$

Tomando $\mathbf{B} = \mathbf{C} = \mathbf{p} + e/c \cdot \mathbf{A}$, y teniendo en cuenta que

$$\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right) \times \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right) = \frac{e}{c}\{\mathbf{p} \times \mathbf{A} + \mathbf{A} \times \mathbf{p}\}$$
$$= -i\hbar e/c \cdot \text{rot } \mathbf{A} = -i\hbar e/c \cdot \mathcal{H},$$

donde $\mathcal{H}$ es el campo magnético, resulta

$$\left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 = \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 + \frac{\hbar e}{c}(\boldsymbol{\sigma},\ \mathcal{H}). \tag{33}$$

Asimismo resulta

$$\left(p_0 + \frac{e}{c}A_0\right)\left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right) - \left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right)\left(p_0 + \frac{e}{c}A_0\right)$$
$$= \frac{e}{c}(\boldsymbol{\sigma},\ p_0\mathbf{A} - \mathbf{A}p_0 + A_0\mathbf{p} - \mathbf{p}A_0)$$
$$= \frac{i\hbar e}{c}\left(\boldsymbol{\sigma}, \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} + \text{grad } A_0\right) = -i\frac{\hbar e}{c}(\boldsymbol{\sigma},\ \mathcal{E}),$$

donde $\mathcal{E}$ es el campo eléctrico. Por lo tanto, (31) da

$$\left\{\left(p_0 + \frac{e}{c}A_0\right)^2 - \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - m^2c^2 - \frac{\hbar e}{c}(\boldsymbol{\sigma},\ \mathcal{H}) + i\rho_1\frac{\hbar e}{c}(\boldsymbol{\sigma},\ \mathcal{E})\right\}\psi = 0. \tag{34}$$

Esta ecuación difiere de (30) en que tiene dos términos más en el operador. Estos términos extra implican ciertos efectos físicos nuevos, pero como aquéllos no son reales no se prestan a una interpretación física directa.

Para comprender los hechos físicos que se derivan de la diferencia entre (34) y (30) es mejor emplear la representación de Heisenberg, que es siempre la más adecuada para establecer comparaciones entre la mecánica clásica y la cuántica. Las ecuaciones de Heisenberg del movimiento están determinadas por el hamiltoniano

$$H = -eA_0 + c\rho_1\left(\sigma,\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right) + \rho_3 mc^2, \tag{35}$$

que se obtiene por generalización de (23) para el caso en que hay campo. La ecuación (35) da, con la ayuda de (33),

$$\left(\frac{H}{c} + \frac{e}{c}A_0\right)^2 = \left\{\rho_1\left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right) + \rho_3 mc\right\}^2$$
$$= \left(\boldsymbol{\sigma},\ \mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 + m^2c^2$$
$$= \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 + m^2c^2 + \frac{\hbar e}{c}(\boldsymbol{\sigma},\ \mathcal{H}) \tag{36}$$

el mismo resultado pues ahora, [illegible] la ecuac
de Dirac [illegible] representación típica, [illegible] que en el
caso no relativista el hamiltoniano de Dirac se reduce al ha
[illegible] de la teoría no relativista, de dos Compo

Aquí aparece la parte real de los términos adicionales de (34) sin su parte imaginaria. Para un electrón que se mueva lentamente (es decir, con momento pequeño), podemos esperar que las ecuaciones del movimiento de Heisenberg estén determinadas por un hamiltoniano de la forma $mc^2 + H_1$, donde H_1 es pequeño en comparación con mc^2. Sustituyendo en (36) H por $mc^2 + H_1$ y despreciando H_1^2 y otros términos que contienen c^{-2}, obtenemos, dividiendo por $2m$

$$H_1 + eA_0 = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\boldsymbol{A}\right)^2 + \frac{\hbar e}{2mc}(\boldsymbol{\sigma}, \boldsymbol{\mathcal{H}}). \tag{37}$$

El hamiltoniano H_1 dado por (37) es el mismo que el hamiltoniano clásico para un electrón lento a excepción del último término

$$\frac{\hbar e}{2mc}(\boldsymbol{\sigma}, \boldsymbol{\mathcal{H}}).$$

Este término puede considerarse como una energía potencial adicional de un electrón lento en la teoría cuántica y puede interpretarse que proviene de que *el electrón tiene un momento magnético* $-\hbar e/2mc \cdot \boldsymbol{\sigma}$. Este momento magnético es el que postulamos en §§ 41 y 47 al estudiar el efecto Zeeman y está de acuerdo con la experiencia.

El momento angular de spin no origina ninguna energía potencial y por ello no aparece en el resultado del cálculo precedente. La manera más simple de mostrar la existencia del momento angular de spin es considerar el caso del movimiento de un electrón libre o de un electrón en un campo de fuerzas central y determinar las integrales del momento angular. Para ello hemos de emplear el hamiltoniano (23), o el hamiltoniano (35) tomando $\mathbf{A} = 0$ y A_0 función del radio r, es decir,

$$H = -eA_0(r) + c\rho_1(\boldsymbol{\sigma}, \mathbf{p}) + \rho_3 mc^2, \tag{38}$$

y obtener las ecuaciones del movimiento de Heisenberg para el momento angular. Con ambos hamiltonianos se obtiene una variación respecto al tiempo de la componente x_1 del momento angular orbital $m_1 = x_2p_3 - x_3p_2$, que vale

$$\begin{aligned} i\hbar\dot{m}_1 &= m_1H - Hm_1 \\ &= c\rho_1\{m_1(\boldsymbol{\sigma}, \mathbf{p}) - (\boldsymbol{\sigma}, \mathbf{p})m_1\} \\ &= c\rho_1(\boldsymbol{\sigma}, m_1\mathbf{p} - \mathbf{p}m_1) \\ &= i\hbar c\rho_1\{\sigma_2p_3 - \sigma_3p_2\}. \end{aligned}$$

en cuya deducción hemos empleado las relaciones de conmutación obtenidas en § 35.

Por lo tanto, $\dot{m}_1 \neq 0$ de donde el momento angular orbital no es una constante del movimiento. Este resultado era de esperar a la vista de la

ecuación del movimiento integrada (29), puesto que la parte oscilatoria del movimiento que aparece en ésta debe dar lugar a un término oscilatorio en el momento angular.

Por otro lado, teniendo en cuenta las ecuaciones (51) de § 37, resulta

$$\begin{aligned} i\hbar\dot{\sigma}_1 &= \sigma_1 H - H\sigma_1 \\ &= c\rho_1\{\sigma_1(\boldsymbol{\sigma}, \mathbf{p}) - (\boldsymbol{\sigma}, \mathbf{p})\sigma_1\} \\ &= c\rho_1(\sigma_1\boldsymbol{\sigma} - \boldsymbol{\sigma}\sigma_1, \mathbf{p}) \\ &= 2ic\rho_1\{\sigma_3 p_2 - \sigma_2 p_3\} \end{aligned}$$

Luego,

$$\dot{m}_1 + \tfrac{1}{2}\hbar\dot{\sigma}_1 = 0,$$

y en consecuencia, el vector $\mathbf{m} + \frac{1}{2}\hbar\boldsymbol{\sigma}$ es una constante del movimiento. Puede interpretarse este resultado diciendo que *el electrón tiene un momento angular de spin* $\frac{1}{2}\hbar\boldsymbol{\sigma}$, que para obtener una constante del movimiento ha de añadirse al momento angular orbital $\mathbf{m}$. El momento angular de spin podía haberse obtenido también a partir de los operadores de rotación para los estados de spin, según el método general de § 35.

El mismo vector $\boldsymbol{\sigma}$ fija las direcciones del momento magnético de spin y del momento angular de spin. Si un electrón en un cierto estado de spin tiene un momento angular de spin $\frac{1}{2}\hbar$ en una dirección particular, tendrá a su vez un momento magnético $-e\hbar/2mc$ en la misma dirección.

Hemos obtenido el valor $\frac{1}{2}\hbar$ para el spin del electrón mediante un argumento basado únicamente en principios generales de la teoría cuántica y de la relatividad. Podríamos aplicar el mismo argumento a otras partículas elementales y llegaríamos a la misma conclusión de que su momento angular de spin es medio cuanto. Esto sería satisfactorio para el protón y el neutrón, pero hay algunas partículas elementales (por ejemplo el fotón y ciertos tipos de mesones) cuyos spines sabemos experimentalmente que son distintos de $\frac{1}{2}\hbar$; he aquí, pues, una discrepancia entre nuestra teoría y la experiencia.

La respuesta ha de encontrarse en una hipótesis que está implícita en nuestro trabajo. El argumento sólo es válido si la posición de la partícula es un observable. Si se verifica esta hipótesis, la partícula ha de tener un momento angular de spin de medio cuanto. Para las partículas que tienen un spin diferente no debe verificarse la hipótesis, y las variables dinámicas x_1, x_2, x_3, que puedan introducirse para describir la posición de la partícula, según nuestra teoría general no serán observables. Para tales partículas no existe una representación de Schrödinger auténtica. Sería posible introducir una cuasi función de onda relativa a las variables dinámicas x_1, x_2, x_3 pero no tendría la interpretación física correcta de una función de onda y el cuadrado de su módulo no representaría, por tanto, la densidad de probabilidad. Para tales partículas existe una representación de momentos que es suficiente para casos prácticos.

(*) [illegible]

71. *Transformación a variables polares*

Para el estudio posterior del movimiento de un electrón en un campo de fuerzas central con el hamiltoniano (38), es conveniente hacer una transformación a coordenadas polares, como hicimos en § 38 en el caso no relativista. Podemos introducir r y p_r igual que allí, pero en vez de k, magnitud del momento angular orbital $\mathbf{m}$, que ya no es una constante del movimiento, hemos de emplear ahora la magnitud del movimiento angular total $\mathbf{M} = \mathbf{m} + \frac{1}{2}\hbar\boldsymbol{\sigma}$. Pongamos

$$j^2\hbar^2 = M_1^2 + M_2^2 + M_3^2 + \tfrac{1}{4}\hbar^2. \tag{39}$$

los autovalores de m_3 son múltiples enteros de $\hbar$, los de $\frac{1}{2}\hbar\sigma^3$ son $\pm\frac{1}{2}\hbar$ y, por consiguiente, los de M_3 han de ser múltiplos impares semienteros de $\hbar$. De la teoría de § 36 se deduce que los autovalores de $|j|$ han de ser enteros positivos.

Si en la fórmula (32) tomamos $\mathbf{B} = \mathbf{C} = \mathbf{m}$, obtenemos

$$\begin{aligned}(\boldsymbol{\sigma}, \mathbf{m})^2 &= \mathbf{m}^2 + i(\boldsymbol{\sigma}, \mathbf{m} \times \mathbf{m}) \\ &= \mathbf{m}^2 - \hbar(\boldsymbol{\sigma}, \mathbf{m}) \\ &= (\mathbf{m} + \tfrac{1}{2}\hbar\boldsymbol{\sigma})^2 - 2\hbar(\boldsymbol{\sigma}, \mathbf{m}) - \tfrac{3}{4}\hbar^2.\end{aligned}$$

De donde

$$\{(\boldsymbol{\sigma}, \mathbf{m}) + \hbar\}^2 = \mathbf{M}^2 + \tfrac{1}{4}\hbar^2.$$

Luego, $(\boldsymbol{\sigma}, \mathbf{m}) + \hbar$ es una cantidad cuyo cuadrado es $\mathbf{M}^2 + \frac{1}{4}\hbar^2$, y de acuerdo con la ecuación (39), podríamos definir $j\hbar$ como $(\boldsymbol{\sigma}, \mathbf{m}) + \hbar$. Ésta no sería, sin embargo, la definición más conveniente de j, pues queremos que j sea una constante del movimiento y $(\boldsymbol{\sigma}, \mathbf{m}) + \hbar$ no lo es. Aplicando (32) tenemos

$$(\boldsymbol{\sigma}, \mathbf{m})(\boldsymbol{\sigma}, \mathbf{p}) = i(\boldsymbol{\sigma}, \mathbf{m} \times \mathbf{p})$$

y

$$(\boldsymbol{\sigma}, \mathbf{p})(\boldsymbol{\sigma}, \mathbf{m}) = i(\boldsymbol{\sigma}, \mathbf{p} \times \mathbf{m}),$$

de modo que

$$\begin{aligned}(\boldsymbol{\sigma}, \mathbf{m})(\boldsymbol{\sigma}, \mathbf{p}) + (\boldsymbol{\sigma}, \mathbf{p})(\boldsymbol{\sigma}, \mathbf{m}) &= i \sum_{123} \sigma_1\{m_2p_3 - m_3p_2 + p_2m_3 - p_3m_2\} \\ &= i \sum_{123} \sigma_1 \cdot 2i\hbar p_1 = -2\hbar(\boldsymbol{\sigma}, \mathbf{p}),\end{aligned}$$

o sea,

$$\{(\boldsymbol{\sigma}, \mathbf{m}) + \hbar\}(\boldsymbol{\sigma}, \mathbf{p}) + (\boldsymbol{\sigma}, \mathbf{p})\{(\boldsymbol{\sigma}, \mathbf{m}) + \hbar\} = 0.$$

Por lo tanto, $(\boldsymbol{\sigma}, \mathbf{m}) + \hbar$ anticonmuta con uno de los términos en la expresión (38) de H, a saber, el término $c\rho_1(\boldsymbol{\sigma}, \mathbf{p})$, y conmuta con los otros dos. De ello se sigue que $\rho_3\{(\boldsymbol{\sigma}, \mathbf{m}) + \hbar\}$ conmuta con los tres términos de H y que, por lo tanto, es una constante del movimiento. Además, el cuadrado

[illegible] Nau de Broglie (1930), Infeld [illegible])

de $\rho_3\{\boldsymbol{\sigma}, \mathbf{m}) + \hbar\}$ es también $\mathbf{M}^2 + \frac{1}{4}\hbar^2$. Podemos tomar, por consiguiente,

$$j\hbar = \rho_3\{(\boldsymbol{\sigma}, \mathbf{m}) + \hbar\}, \tag{40}$$

que es una definición racional y conveniente de j, en concordancia con (39), y hace de j una constante del movimiento. Los autovalores de esta j, son todos los enteros positivos y negativos menos el cero.

Mediante una nueva aplicación de (32) obtenemos, con la ayuda de (40) y también de la ecuación (58) de § 38

$$\begin{aligned}(\boldsymbol{\sigma}, \mathbf{x})(\boldsymbol{\sigma}, \mathbf{p}) &= (\mathbf{x}, \mathbf{p}) + i(\boldsymbol{\sigma}, \mathbf{m}) \\ &= rp_r + i\rho_3 j\hbar - i\hbar,\end{aligned} \tag{41}$$

Introducimos el operador lineal ε definido por

$$r\varepsilon = \rho_1(\boldsymbol{\sigma}, \mathbf{x}). \tag{42}$$

Como r conmuta con ρ_1 y con $(\boldsymbol{\sigma}, \mathbf{x})$ ha de conmutar con ε. Tenemos pues

$$r^2\varepsilon^2 = [\rho_1(\boldsymbol{\sigma}, \mathbf{x})]^2 = (\boldsymbol{\sigma}, \mathbf{x})^2 = \mathbf{x}^2 = r^2,$$

o sea,

$$\varepsilon^2 = 1.$$

Ahora bien, $\rho_1(\boldsymbol{\sigma}, \mathbf{p})$ conmuta con j y como en lo que respecta al momento angular hay simetría entre $\mathbf{x}$ y $\mathbf{p}$, $\rho_1(\boldsymbol{\sigma}, \mathbf{x})$ ha de conmutar también con j. Luego, ε conmuta con j. Además ε ha de conmutar con p_r, puesto que tenemos

$$(\boldsymbol{\sigma}, \mathbf{x})(\mathbf{x}, \mathbf{p}) - (\mathbf{x}, \mathbf{p})(\boldsymbol{\sigma}, \mathbf{x}) = (\boldsymbol{\sigma}, \mathbf{x}(\mathbf{x}, \mathbf{p}) - (\mathbf{x}, \mathbf{p})\mathbf{x}) = i\hbar(\boldsymbol{\sigma}, \mathbf{x}),$$

que da

$$r\varepsilon rp_r - rp_r r\varepsilon = i\hbar r\varepsilon,$$

o bien

$$r^2\varepsilon p_r - r^2 p_r\varepsilon = 0.$$

De (41) y (42) obtenemos

$$r\varepsilon\rho_1(\boldsymbol{\sigma}, \mathbf{p}) = rp_r + i\rho_3 j\hbar - i\hbar,$$

o sea,

$$\rho_1(\boldsymbol{\sigma}, \mathbf{p}) = \varepsilon(p_r - i\hbar/r) + i\varepsilon\rho_3 j\hbar/r.$$

Con lo que (38) da

$$H/c = -e/c \cdot A_0 + \varepsilon(p_r - i\hbar/r) + i\varepsilon\rho_3 j\hbar/r + \rho_3 mc.$$

que expresa el hamiltoniano en función de las variables polares. Debemos observar que ε y ρ_3 conmutan con todas las demás variables que figuran en H, y anticonmutan entre sí. Esto significa que podemos tomar una

representación con ρ_3 diagonal en la cual ε y ρ_3 están representados respectivamente por las matrices

$$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{43}$$

Si r es también diagonal en la representación, el representante $\langle r'\rho_3'|\rangle$ de un ket tendrá dos componentes $\langle r', 1|\rangle = \psi_a(r')$ y $\langle r', -1|\rangle = \psi_b(r')$, relativas a las dos filas y columnas de las matrices (43).

72. *Estructura fina de los niveles de energía del hidrógeno*

Consideremos ahora el caso del átomo de hidrógeno, para el cual $A_0 = e/r$, y calculemos sus niveles de energía dados por los autovalores H' de H. La ecuación $(H' - H)|\rangle = 0$ que define dichos autovalores, escrita en función de sus representantes en la representación considerada anteriormente en la que ε y ε_3 están representados por las matrices (43), da lugar a las ecuaciones

$$\left(\frac{H'}{c} + \frac{e^2}{cr}\right)\psi_a + \hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)\psi_b + \frac{j\hbar}{r}\psi_b - mc\psi_a = 0,$$
$$\left(\frac{H'}{c} + \frac{e^2}{cr}\right)\psi_b - \hbar\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)\psi_a + \frac{j\hbar}{r}\psi_a + mc\psi_b = 0.$$

Si ponemos

$$\frac{\hbar}{mc - H'/c} = a_1, \qquad \frac{\hbar}{mc + H'/c} = a_2, \tag{44}$$

estas ecuaciones se reducen a

$$\left.\begin{aligned} \left(\frac{1}{a_1} - \frac{\alpha}{r}\right)\psi_a - \left(\frac{\partial}{\partial r} + \frac{j+1}{r}\right)\psi_b &= 0, \\ \left(\frac{1}{a_2} + \frac{\alpha}{r}\right)\psi_b - \left(\frac{\partial}{\partial r} - \frac{j-1}{r}\right)\psi_a &= 0, \end{aligned}\right\} \tag{45}$$

donde $\alpha = e^2/\hbar c$, que es un número pequeño. Vamos a resolver estas ecuaciones por un método semejante al que empleábamos para la ecuación (73) de § 39.

Pongamos

$$\psi_a = r^{-1}e^{-r/a}f, \qquad \psi_b = r^{-1}e^{-r/a}g, \tag{46}$$

introduciendo las dos nuevas funciones de r, f y g, siendo

$$a = (a_1a_2)^{\frac{1}{2}} = \hbar(m^2c^2 - H'^2/c^2)^{-\frac{1}{2}}. \tag{47}$$

Las ecuaciones (45) dan

$$\left.\begin{aligned}\left(\frac{1}{a_1}-\frac{\alpha}{r}\right)f-\left(\frac{\partial}{\partial r}-\frac{1}{a}+\frac{j}{r}\right)g &= 0,\\ \left(\frac{1}{a_2}+\frac{\alpha}{r}\right)g-\left(\frac{\partial}{\partial r}-\frac{1}{a}-\frac{j}{r}\right)f &= 0.\end{aligned}\right\} \tag{48}$$

Vamos a ensayar una solución para f y g de la forma de una serie de potencias

$$f = \sum_s c_s r^s, \qquad g = \sum_s c'_s r^s, \tag{49}$$

donde los valores consecutivos de s difieren en una unidad si bien no exigimos que sean enteros. Sustituyendo estas expresiones de f y g en (48) y separando los coeficientes de r^{s-1}, obtenemos

$$\left.\begin{aligned}c_{s-1}/a_1 - \alpha c_s - (s+j)c'_s + c'_{s-1}/a &= 0,\\ c'_{s-1}/a_2 + \alpha c'_s - (s-j)c_s + c_{s-1}/a &= 0.\end{aligned}\right\} \tag{50}$$

Multiplicando la primera de estas ecuaciones por a y la segunda por a_2 y restando, eliminamos c_{s-1} y c'_{s-1}, puesto que según (47) $a/a_1 = a_2/a$. Nos queda

$$[a\alpha - a_2(s-j)]c_s + [a_2\alpha + a(s+j)]c'_s = 0, \tag{51}$$

relación que liga las c y las c'.

La condición límite para $r = 0$ exige que $r\psi_a$ y $r\psi_b$ tiendan a cero cuando $r \to 0$, luego, según (46), f y g tienden a cero para $r \to 0$. Por lo tanto, las series (49) han de terminar por el lado de las s pequeñas. Si s_0 es el valor mínimo de s para el cual c_s y c'_s no son ambas nulas, de (50) resulta, poniendo $s = s_0$ y $c_{s_0-1} = c'_{s_0-1} = 0$,

$$\left.\begin{aligned}\alpha c_{s_0} + (s_0+j)c'_{s_0} &= 0,\\ \alpha c'_{s_0} - (s_0-j)c_{s_0} &= 0,\end{aligned}\right\} \tag{52}$$

que da

$$\alpha^2 = -s_0^2 + j^2.$$

Como la condición límite exige que el valor mínimo de s sea mayor que cero, hemos de tomar

$$s_0 = +\sqrt{(j^2-\alpha^2)}.$$

Para estudiar la convergencia de las series (49) hemos de hallar la relación c_s/c_{s-1} para s grande. La ecuación (51) y la segunda ecuación (50) da aproximadamente, cuando s es grande,

$$a_2 c_s = a c'_s$$

y

$$sc_s = c_{s-1}/a + c'_{s-1}/a_2.$$

De donde

$$c_s/c_{s-1} = 2/as.$$

Las series (49) convergerán, por lo tanto, como

$$\sum_s \frac{1}{s!}\left(\frac{2r}{a}\right)^s,$$

o $e^{2r/a}$. Este resultado es análogo al obtenido en § 39 y nos permite inferir, igual que allí, que para a imaginario puro, o sea, según (47), para $H' > mc^2$, son posibles todos los valores de H', mientras que para $H' < mc^2$, eligiendo a positivo, resulta entonces que sólo son posibles los valores de H' para los cuales las series (49) terminan por el lado de las s grandes.

Si las series (49) terminan con los términos c_s y c'_s o sea si $c_{s+1} = c'_{s+1} = 0$, sustituyendo s por $s+1$ en (50) obtendremos

$$\left.\begin{aligned} c_s/a_1 + c'_s/a &= 0, \\ c'_s/a_2 + c_s/a &= 0. \end{aligned}\right\} \quad (53)$$

Estas dos ecuaciones según (47) son equivalentes. Combinadas con (51) dan

$$a_1[a\alpha - a_2(s-j)] = a[a_2\alpha + a(s+j)],$$

que se reduce a

$$2a_1a_2s = a(a_1 - a_2)\alpha,$$

o con la ayuda de (44),

$$\frac{s}{a} = \frac{1}{2}\left(\frac{1}{a_2} - \frac{1}{a_1}\right)\alpha = \frac{H'}{c\hbar}\alpha,$$

Elevando al cuadrado y empleando (47) obtenemos

$$s^2(m^2c^2 - H'^2/c^2) = \alpha^2H'^2/c^2.$$

De donde

$$\frac{H'}{mc^2} = \left(1 + \frac{\alpha^2}{s^2}\right)^{-\frac{1}{2}}.$$

La s de esta ecuación, que determina el último término de las series, ha de superar a s_0 en algún entero no menor que cero. Llamando n a dicho entero, tenemos

$$s = n + \sqrt{(j^2 - \alpha^2)}$$

y por consiguiente,

$$\frac{H'}{mc^2} = \left\{ 1 + \frac{\alpha^2}{\{n + \sqrt{(j^2 - \alpha^2)}\}^2} \right\}^{-\frac{1}{2}}. \tag{54}$$

Esta fórmula nos da los niveles de energía discretos del espectro del hidrógeno, y fue obtenida por primera vez por Sommerfeld trabajando con la teoría de las órbitas de Bohr. Contiene dos números cuánticos n y j, pero, por ser α^2 muy pequeño, la energía depende casi exclusivamente de $n + |j|$. Los valores de n y de $|j|$ que dan el mismo valor de $n + |j|$, dan lugar a un conjunto de niveles de energía situados muy cerca unos de otros y, salvo el término constante mc^2, también muy cerca del nivel de energía dado por la fórmula no relativista (80) de § 39 para $s = n + |j|$.

Hemos utilizado las ecuaciones (53) combinadas con (51), con lo cual no hemos empleado a fondo (53), puesto que los coeficientes de c_s y c'_s en (51) pueden ser ambos nulos. En este caso, multiplicando el primer coeficiente por a_1 y el segundo por a y sumando, obtenemos

$$a(a_1 + a_2)\alpha + 2a_1a_2j = 0.$$

Luego, j ha de ser negativo en este caso. Con ayuda de (44) y (47) obtenemos además

$$-\frac{2j}{\alpha} = \frac{a}{a_2} + \frac{a}{a_1} = \frac{2mca}{\hbar} = \frac{2mc}{(m^2c^2 - H'^2/c^2)^{\frac{1}{2}}}$$

o sea,

$$\frac{H'^2}{m^2c^4} = 1 - \frac{\alpha^2}{j^2}$$

Como H' ha de ser positivo, llegamos a

$$\frac{H'}{mc^2} = \frac{\sqrt{(j^2 - \alpha^2)}}{|j|} \tag{55}$$

que es el valor de H' dado por (54) para $n = 0$. El caso $n = 0$ y j negativo necesita, por lo tanto, otro examen para ver si se cumplen las condiciones (53).

Con $n = 0$ el valor máximo de s es igual al mínimo, luego, las ecuaciones (53) con s_0 en lugar de s tienen que concordar con (52). Ahora bien, aplicando (44) y (47), (55) da

$$\frac{1}{a_1} = \frac{mc}{\hbar}\left(1 - \frac{\sqrt{(j^2 - \alpha^2)}}{|j|}\right) \qquad \frac{1}{a} = \frac{mc}{\hbar}\frac{\alpha}{|j|}$$

luego, la primera ecuación (53), poniendo s_0 en lugar de s, da

$$c_{s_0}\{|j| - \sqrt{(j^2 - \alpha^2)}\} + c'_{s_0}\,\alpha = 0,$$

que concuerda con la segunda ecuación (52) solamente si j es positivo. Concluimos, pues que para $n = 0$, j ha de ser entero positivo, mientras que para los otros valores de n, j puede tomar todos los valores enteros distintos de cero.

73. *Teoría del positrón*

En § 67 decíamos que la ecuación de onda para el electrón admite doble número de soluciones de las que debería admitir; la mitad de ellas correspondían a estados con valores negativos de la energía cinética $cp_0 + eA_0$. Esta dificultad se introdujo desde el momento en que pasamos de la ecuación (5) a la ecuación (6), y es inherente a toda teoría relativista. Aparece también en teoría relativista clásica, pero allí no constituye una dificultad seria debido a que todas las variables dinámicas clásicas varían con continuidad, y si la energía cinética $cp_0 + eA_0$ es inicialmente positiva (en cuyo caso tiene que ser mayor o igual que mc^2) no puede ser negativa en un instante posterior (en cuyo caso tendría que ser menor o igual que $-mc^2$). En cambio, en la teoría cuántica pueden presentarse transiciones discontinuas, y si el electrón están inicialmente en un estado de energía cinética positiva puede hacer una transición y pasar a un estado de energía cinética negativa. Ya no es posible, por lo tanto, ignorar sencillamente los estados de energía negativa como en teoría clásica.

Examinemos más detenidamente las soluciones de energía negativa de la ecuación

$$\left\{\left(p_0 + \frac{e}{c}A_0\right) - \alpha_1\left(p_1 + \frac{e}{c}A_1\right) - \right.$$
$$\left. - \alpha_2\left(p_2 + \frac{e}{c}A_2\right) - \alpha_3\left(p_3 + \frac{e}{c}A_3\right) - \alpha_m mc\right\}\psi = 0 \qquad (56)$$

Para ello es conveniente usar una representación de las α en la que todos los elementos de las matrices que representan α_1, α_2 y α_3 sean reales y todos los que representan α_m sean imaginarios puros o cero. Puede obtenerse una representación de este tipo, por ejemplo, a partir de la de § 67 intercambiando las expresiones de α_2 y α_m en (9). Si expresamos (56) como una ecuación matricial en esta representación y ponemos $-i$ en lugar de i, recordando la i de (4), obtendremos

$$\left\{\left(-p_0 + \frac{e}{c}A_0\right) - \alpha_1\left(-p_1 + \frac{e}{c}A_1\right) - \right.$$
$$\left. - \alpha_2\left(-p_2 + \frac{e}{c}A_2\right) - \alpha_3\left(-p_3 + \frac{e}{c}A_3\right) + \alpha_m mc\right\}\bar{\psi} = 0. \qquad (57)$$

Por lo tanto, cada solución ψ de la ecuación de ondas (56) tiene por compleja conjugada $\bar{\psi}$ una solución de la ecuación de ondas (57). Además si la solución ψ de (56) pertenece a un valor negativo de $cp_0 + eA_0$ la correspondiente solución $\bar{\psi}$ de (57) pertenecerá a un valor positivo de $cp_0 + eA_0$. Pero el operador de (57) es precisamente el que obtendría sustituyendo e por $-e$ en el operador de (56). De ello se deduce que cada solución de energía negativa de (56) es la compleja conjugada de una solución de energía positiva de la ecuación de onda que se obtiene de (56) sustituyendo e por $-e$, solución que representa un electrón de carga $+e$ (en vez de $-e$ como teníamos hasta ahora) moviéndose a través del campo electromagnético dado. Por lo tanto, las soluciones extra de (56) están relacionadas con el movimiento de un electrón de carga $+e$. (Es imposible separar definidamente las soluciones de (56) para un campo electromagnético general en soluciones referidas a valores positivos y a valores negativos de $cp_0 + eA_0$, puesto que tal separación implicaría que no pudieran tener lugar transiciones de un tipo de soluciones a otro. La discusión precedente es solamente aproximada, y sólo se aplica cuando es aproximadamente posible tal separación.)

Deducimos así que las soluciones de (56) que pertenecen a una energía negativa se refieren al movimiento de un nuevo tipo de partícula de masa igual a la del electrón y de carga opuesta. Tales partículas se han observado experimentalmente y se llaman *positrones*. Sin embargo, no podemos afirmar simplemente que las soluciones de energía negativa representen positrones, puesto que con ello todas las relaciones dinámicas serían incorrectas. Por ejemplo, es evidente que un positrón no tiene una energía cinética negativa. Así pues, hemos de establecer la teoría del positrón sobre una base algo distinta. Supongamos que casi todos los estados de energía negativa están ocupados, con un electrón en cada estado de acuerdo con el principio de exclusión de Pauli. Un estado de energía negativa desocupado, aparecerá entonces como algo con energía positiva, puesto que para hacerlo desaparecer, es decir, para llenarlo, tendríamos que añadir un electrón de energía negativa. Hacemos la hipótesis de que *estos estados desocupados de energía negativa son los positrones.*

Estas hipótesis exigen que haya una distribución de electrones de densidad infinita en cualquier lugar del espacio. Un vacío perfecto es una región donde todos los estados de energía positiva están desocupados y todos los de energía negativa están ocupados. Por supuesto, en un vacío perfecto ha de ser válida la ecuación de Maxwell

$$\operatorname{div} \mathcal{E} = 0.$$

Esto significa que la distribución infinita de electrones de energía nega-

tiva no contribuye al campo electrónico. Únicamente contribuirán a la densidad eléctrica j_0 de la ecuación de Maxwell

$$\operatorname{div} \mathcal{E} = 4\pi j_0. \tag{58}$$

las desviaciones respecto a la distribución de vacío perfecto.

Habrá, por lo tanto, una contribución $-e$ por cada estado de energía positiva ocupado y una contribución $+e$ por cada estado de energía negativa desocupado.

El principio de exclusión actuará corrientemente para impedir que un electrón de energía positiva haga transiciones a estados de energía negativa. Sin embargo, siempre será posible para este electrón pasar a un estado de energía negativa desocupado. En este caso desaparecerán simultáneamente un electrón y un positrón, emitiéndose su energía en forma de radiación. El proceso inverso consistiría en la creación de un electrón y un positrón a partir de radiación electromagnética.

De la simetría entre estados fermiónicos ocupados y desocupados discutida al final de § 65, deducimos que la presente teoría es simétrica respecto a los electrones y positrones. Si supusiéramos que los positrones fuesen las partículas básicas descritas por ecuaciones de onda de la forma (11) con $-e$ en vez de e, suponiendo que casi todos los estados de energía negativa de los positrones están ocupados, e interpretando un hueco en la distribución de positrones de energía negativa como un electrón ordinario, tendríamos una teoría equivalente. Dicha teoría podría desarrollarse de acuerdo con la hipótesis de que todas las leyes de la física son simétricas respecto a cargas eléctricas positivas y negativas.

XII

ELECTRODINÁMICA CUÁNTICA

74. *El campo electromagnético en ausencia de materia*

En la teoría de la radiación establecida en el capítulo X, al estudiar la interacción de la radiación con la materia, nos vimos obligados a introducir algunas aproximaciones. El objeto del presente capítulo es eliminar estas aproximaciones y obtener, dentro de lo posible, una teoría rigurosa de la interacción del campo electromagnético con la materia, sujeta a la limitación de que la materia conste solamente de electrones y positrones. Las otras formas de materia (protones, neutrones, etc.), son todavía poco conocidas para intentar obtener de momento una teoría rigurosa de su interacción con el campo electromagnético. En cambio, existe una buena teoría precisa de los electrones y positrones, dada en el capítulo anterior, que permite construir una teoría precisa de la interacción del campo electromagnético con dichas partículas. La teoría ha de contener tanto la interacción de los electrones y positrones entre sí, dada por las fuerzas de Coulomb, como su interacción con la radiación electromagnética, y ha de satisfacer, por supuesto, los principios de la relatividad restringida. Para abreviar tomaremos en este capítulo $c = 1$.

En primer lugar consideraremos el campo electromagnético sin interacción con la materia. En § 63 establecimos por primera vez un modo de tratar el campo de radiación sin interacción con la materia. Introdujimos variables dinámicas para describir el campo, establecimos las relaciones de conmutación entre ellas, y encontramos un hamiltoniano que explicaba correctamente su variación con el tiempo. En esta ocasión no hicimos aproximaciones. Por lo tanto, la teoría resultante sería exacta y satisfactoria si no fuera por haber tomado el potencial escalar igual a cero. Esto destruye la forma relativista de la teoría y la hace inadecuada como punto de partida para desarrollar una teoría rigurosa del campo electromagnético en interacción con la materia.

Tenemos que extender, por lo tanto, el tratamiento de § 63 conservando un A_0 general e introduciéndolo en la teoría junto con los otros

potenciales A_1, A_2, A_3. Las cuatro A_μ que así tendremos, verificarán las ecuaciones que se obtienen generalizando (62) de § 63,

$$\Box A_\mu = 0, \tag{1}$$

$$\partial A_\mu / \partial x_\mu = 0. \tag{2}$$

Por ahora dejaremos de lado la segunda de estas ecuaciones y solamente consideraremos la primera.

La ecuación (1) muestra que cada A_μ puede descomponerse en ondas que avanzan con la velocidad de la luz. Así pues, en lugar de la ecuación (63) de § 63 resultará ahora

$$A_\mu(x) = \int (A_{\mu\mathbf{k}} e^{ik.x} + \bar{A}_{\mu\mathbf{k}} e^{-ik.x})\, d^3k, \tag{3}$$

donde $k \cdot x$ indica el producto escalar cuatridimensional

$$k \cdot x = k_0 x_0 - (\mathbf{k}, \mathbf{x}),$$

siendo k_ν el cuadrivector cuyas componentes espaciales son iguales a las componentes del vector tridimensional $\mathbf{k}$ de § 63 y cuya componente temporal es $k_0 = |\mathbf{k}|$; d^3k indica $dk_1 dk_2 dk_3$, como en § 63.

La componente de Fourier $A_{\mu\mathbf{k}}$ tiene una parte $A_{0\mathbf{k}}$ procedente de $A_0(x)$ y una parte $A_{r\mathbf{k}}$ ($r = 1, 2, 3$), que es un vector tridimensional. Este último puede descomponerse en dos partes, una parte longitudinal dirigida según la dirección de $\mathbf{k}$, que es la dirección de movimiento de las ondas, y una parte transversal perpendicular a $\mathbf{k}$. La parte longitudinal es $k_r k_s / k_0^2 \cdot A_{s\mathbf{k}}$. La parte transversal es

$$(\delta_{rs} - k_r k_s / k_0^2) A_{s\mathbf{k}} = \mathscr{A}_{r\mathbf{k}}, \tag{4}$$

que verifica la condición

$$k_r \mathscr{A}_{r\mathbf{k}} = 0. \tag{5}$$

Se sabe por la teoría de Maxwell de la luz que solamente la parte transversal actúa dando radiación electromagnética. En el capítulo X únicamente se consideró la parte transversal: la $A_{r\mathbf{k}}$ de § 63 es la misma que la $\mathscr{A}_{r\mathbf{k}}$ de ahora y la ecuación (65) de § 63 corresponde a la ecuación (5). Sin embargo, la parte longitudinal no puede despreciarse en una teoría completa de la electrodinámica a causa de su relación con las fuerzas de Coulomb, como veremos más adelante.

Podemos descomponer ahora el vector tridimensional $A_r(x)$ en dos partes, una parte transversal con sólo componentes de Fourier transversales y una parte longitudinal con sólo componentes de Fourier longitudinales. La primera es

$$\mathscr{A}_r(x) = \int (\mathscr{A}_{r\mathbf{k}}\, e^{ik\cdot x} + \bar{\mathscr{A}}_{r\mathbf{k}}\, e^{-ik\cdot x})\, d^3k$$

que verifica

$$\partial \mathscr{A}_r(x)/\partial x_r = 0. \tag{6}$$

La parte longitudinal puede expresarse como el gradiente $\partial V/\partial x_r$ de un escalar dado por

$$V = i \int k_s/k_0^2 \cdot (A_{s\mathbf{k}}\, e^{ik\cdot x} - \bar{A}_{s\mathbf{k}}\, e^{-ik\cdot x})\, d^3\mathbf{k}. \tag{7}$$

Luego,

$$A_r = \mathscr{A}_r + \partial V/\partial x_r. \tag{8}$$

El campo magnético queda determinado por la parte transversal de A_r

$$\mathscr{H} = \text{rot } \mathbf{A} = \text{rot } \mathscr{A}.$$

Es conveniente tomar $A_0(x)$ longitudinal, y así los potenciales completos $A_\mu(x)$ quedan separados en una parte transversal $\mathscr{A}_r(x)$ y una parte longitudinal A_0, $\partial V/\partial x_r$. Esta separación se refiere, por supuesto, a un sistema de referencia de Lorentz particular y no ha de usarse cuando se quiere conservar una ecuación en forma relativista.

Cada coeficiente de Fourier $A_{\mu\mathbf{k}}$ aparece en (3) combinada con el factor temporal $e^{ik_0x_0}$. El producto

$$A_{\mu\mathbf{k}}\, e^{ik_0x_0} = A_{\mu\mathbf{k}t}, \tag{9}$$

es una variable dinámica canónica en mecánica clásica, y en mecánica cuántica es una variable dinámica de Heisenberg. Hemos de obtener ahora los P.B. entre estas variables.

El desarrollo de § 63 nos da los P.B. para la parte transversal de $A_{\mu\mathbf{k}t}$. Para determinarlos, pasamos a valores discretos de $\mathbf{k}$ en el espacio tridimensional de $\mathbf{k}$ y tomamos un valor discreto de $\mathbf{k}$ particular, por ejemplo, $k_1 = k_2 = 0$, $k_3 = k_0 > 0$. La variable de polarización $\mathbf{l}$ puede tomar entonces dos valores relativos a las dos direcciones 1 y 2, y la ecuación (73) de § 63, con la ayuda de las relaciones de conmutación de las $\dot{\eta}$ y las $\bar{\eta}$, o sea las ecuaciones (11) de § 60, da

$$[\bar{A}_{1\mathbf{k}t}, A_{1\mathbf{k}t}] = [\bar{A}_{2\mathbf{k}t}, A_{2\mathbf{k}t}] = -is_{\mathbf{k}}/4\pi^2k_0. \tag{10}$$

El desarrollo de § 63 no nos da información sobre $A_{3\mathbf{k}t}$ y $A_{0\mathbf{k}t}$.

Sin embargo, podemos obtener ahora los P.B. para $A_{3\mathbf{k}t}$ y $A_{0\mathbf{k}t}$ a partir de la teoría de la relatividad. Las ecuaciones (10) han de completarse para formar un conjunto relativista, y la única manera sencilla de hacerlo es añadiendo las dos ecuaciones

$$[\bar{A}_{3\mathbf{k}t}, A_{3\mathbf{k}t}] = -[\bar{A}_{0\mathbf{k}t}, A_{0\mathbf{k}t}] = -is_{\mathbf{k}}/4\pi^2k_0, \tag{11}$$

de modo que las cuatro ecuaciones (10) y (11), junto con las condiciones de que $\bar{A}_{\mu\mathbf{k}t}$ y $A_{\nu\mathbf{k}t}$ conmuten para $\mu \neq \nu$ (ya que se refieren a grados de li-

bertad distintos), se puedan reunir para formar la ecuación tensorial única

$$[\bar{A}_{\mu\mathbf{k}t}, A_{\nu\mathbf{k}t}] = ig_{\mu\nu}s_{\mathbf{k}}/4\pi^2k_0. \tag{12}$$

Obtenemos así los P.B., para todas las variables dinámicas. La ecuación (12) puede generalizarse así

$$[\bar{A}_{\mu\mathbf{k}t}, A_{\nu\mathbf{k}'t}] = ig_{\mu\nu}s_{\mathbf{k}}\delta_{\mathbf{k}\mathbf{k}'}/4\pi^2k_0. \tag{13}$$

Consideremos de nuevo valores continuos de $\mathbf{k}$. Para transformar $\delta_{\mathbf{k}\mathbf{k}'}$ a valores continuos de $\mathbf{k}$ observemos que, para una función general $f(\mathbf{k})$ en un espacio tridimensional de $\mathbf{k}$

$$\sum_{\mathbf{k}} f(\mathbf{k})\delta_{\mathbf{k}\mathbf{k}'} = f(\mathbf{k}') = \int f(\mathbf{k})\delta(\mathbf{k}-\mathbf{k}')d^3k, \tag{14}$$

donde $\delta(\mathbf{k}-\mathbf{k}')$ es la función δ tridimensional

$$\delta(\mathbf{k}-\mathbf{k}') = \delta(k_1-k_1')\delta(k_2-k_2')\delta(k_3-k_3').$$

Para que (14) coincida con la ecuación tipo que relaciona sumas e integrales, es decir, la ecuación (52) de § 62, debe ser

$$s_{\mathbf{k}}\delta_{\mathbf{k}\mathbf{k}'} = \delta(\mathbf{k}-\mathbf{k}'). \tag{15}$$

y así (13) da

$$[\bar{A}_{\mu\mathbf{k}t}, A_{\nu\mathbf{k}'t}] = ig_{\mu\nu}/4\pi^2k_0 \cdot \delta(\mathbf{k}-\mathbf{k}'). \tag{16}$$

Esta ecuación, junto con las ecuaciones

$$[A_{\mu\mathbf{k}t}, A_{\nu\mathbf{k}'t}] = [\bar{A}_{\mu\mathbf{k}t}, \bar{A}_{\nu\mathbf{k}'t}] = 0, \tag{17}$$

constituyen los P.B. en la teoría de valores continuos de $\mathbf{k}$. Debemos tener en cuenta que estos P.B. conservan el mismo valor si se omite en ellos el subíndice t refiriéndose entonces a los coeficientes de Fourier constantes $\bar{A}_{\mu\mathbf{k}}$, $A_{\nu\mathbf{k}}$.

Hemos de obtener ahora un hamiltoniano que haga variar en la imagen de Heisenberg cada variable dinámica $A_{\mu\mathbf{k}t}$ con el tiempo $t = x_0$, de acuerdo con la ley (9) con $A_{\mu\mathbf{k}}$ constante. Llamando H_F a este hamiltoniano, debe ser

$$[A_{\mu\mathbf{k}t}, H_F] = dA_{\mu\mathbf{k}t}/dx_0 = ik_0A_{\mu\mathbf{k}t}. \tag{18}$$

Puede verse fácilmente que esta ecuación es satisfecha por

$$H_F = -4\pi^2 \int k_0{}^2A_{\mu\mathbf{k}t}\bar{A}^{\mu}{}_{\mathbf{k}t}d^3k. \tag{19}$$

Tomamos, por lo tanto, (19), con la posible adición de un término numérico arbitrario que no contenga ninguna variable dinámica, como el hamiltoniano del campo electromagnético en ausencia de materia.

En § 63 empleábamos nuestro conocimiento de la parte tranversal del

hamiltoniano para obtener los P.B. de las variables transversales. Después hemos aplicado el procedimiento inverso a las variables longitudinales: hemos empleado nuestro conocimiento de los P.B., obtenidos mediante un argumento relativista, para hallar la parte del hamiltoniano que hace referencia a ellas, eligiéndola de forma que esté de acuerdo con (18).

Si desarrollamos el hamiltoniano (19) obtenemos

$$H_F = 4\pi^2 \int k_0^2(A_{1\mathbf{k}t}\bar{A}_{1\mathbf{k}t} + A_{2\mathbf{k}t}\bar{A}_{2\mathbf{k}t} + A_{3\mathbf{k}t}\bar{A}_{3\mathbf{k}t} - A_{0\mathbf{k}t}\bar{A}_{0\mathbf{k}t})\, d^3k.$$

Los primeros tres términos del integrando tienen una parte transversal que es igual a la energía transversal dada por (71) de § 63. El último término del integrando, que es la parte de H_F relativa al potencial escalar A_0, aparece con un signo menos. Este signo menos está determinado por la relatividad y nos dice que el sistema dinámico formado por las variables $A_{0\mathbf{k}t}$, $\bar{A}_{0\mathbf{k}t}$ es un *oscilador armónico de energía negativa.* Es bastante sorprendente que aparezca así en la teoría un concepto tan poco físico como es el de energía negativa. Veremos en § 77 que la energía negativa asociada a los grados de libertad que provienen de A_0 siempre queda compensada por la energía positiva asociada a los otros grados de libertad longitudinales, de modo que en la práctica nunca se manifiesta.

75. *Expresión relativista de las condiciones cuánticas*

Las ecuaciones de campo de la teoría expuesta en la sección precedente eran relativistas. Para que la teoría sea completamente relativista debemos demostrar que también son relativistas los P.B. En su forma (16) expresada en función de componentes de Fourier la cuestión no es evidente. Vamos a obtener una expresión relativista de los P.B. desarrollando $[A_\mu(x), A_\nu(x')]$, siendo x y x' dos puntos del espacio-tiempo. Pero para ello hemos de estudiar previamente una función del espacio-tiempo singular e invariante.

La función $\delta(x_\mu x^\mu)$ es evidentemente invariante frente a las transformaciones de Lorentz. Se anula en todo el espacio excepto sobre el cono de luz que tiene como vértice el origen, es decir, sobre el espacio tridimensional $x_\mu x^\mu = 0$. Dicho cono de luz consta de dos partes diferentes, una *parte de futuro* para la cual $x_0 > 0$, y una *parte de pasado* para la cual $x_0 < 0$. La función que sobre la parte de futuro del cono de luz toma el valor $\delta(x_\mu x^\mu)$ y sobre la parte de pasado el valor $-\delta(x_\mu x^\mu)$, también es invariante frente a las transformaciones de Lorentz. Dicha función, que puede escribirse en la forma $\delta(x_\mu x^\mu)x_0/|x^0|$, juega un papel importante en la teoría dinámica de campos, por lo cual introducimos una notación especial para designarla. Por definición,

$$\Delta(x) = 2\delta(x_\mu x^\mu)x_0/|x_0|. \tag{20}$$

Esta definición da sentido a la función Δ aplicada a cualquier cuadrivector. Con ayuda de (9) de § 15, podemos expresar $\delta(x_\mu x^\mu)$ en la forma

$$\delta(x_\mu x^\mu) = \tfrac{1}{2}|\mathbf{x}|^{-1}\{\delta(x_0 - |\mathbf{x}|) + \delta(x_0 + |\mathbf{x}|)\}, \tag{21}$$

donde $|\mathbf{x}|$ es la longitud de la parte tridimensional de x_μ; con ello $\Delta(x)$ toma la forma

$$\Delta(x) = |\mathbf{x}|^{-1}\{\delta(x_0 - |\mathbf{x}|) - \delta(x_0 + |\mathbf{x}|)\}. \tag{22}$$

El valor de $\Delta(x)$ en el origen se toma igual a cero, y en consecuencia, $\Delta(-x) = -\Delta(x)$.

Hagamos un análisis de Fourier de $\Delta(x)$. Poniendo d^4x en lugar de $dx_0dx_1dx_2dx_3$, y d^3x en lugar de $dx_1dx_2dx_3$ para un cuadrivector k_μ cualquiera se tiene,

$$\begin{aligned}\int \Delta(x)e^{ik.x}\,d^4x &= \int |\mathbf{x}|^{-1}\{\delta(x_0 - |\mathbf{x}|) - \delta(x_0 + |\mathbf{x}|)\}e^{[ik_0x_0 - (\mathbf{k}\mathbf{x})]}\,d^4x \\ &= \int |\mathbf{x}|^{-1}\{e^{ik_0|\mathbf{x}|} - e^{-ik_0|\mathbf{x}|}\}e^{-i(\mathbf{k}\mathbf{x})}\,d^3x.\end{aligned}$$

Introduciendo ahora coordenadas polares $|\mathbf{x}|$, θ, ϕ en el espacio tridimensional $x_1x_2x_3$ y tomando como polo la dirección de la parte tridimensional *de* k_μ, resulta

$$\begin{aligned}\int \Delta(x)e^{ik\cdot x}\,d^4x &= \iiint \{e^{ik_0|\mathbf{x}|} - e^{-ik_0|\mathbf{x}|}\}\,e^{-|\mathbf{k}||\mathbf{x}|\cos\theta}|\mathbf{x}|\operatorname{sen}\theta d\theta d\phi d|\mathbf{x}| \\ &= 2\pi\int_0^\infty \{e^{ik_0|\mathbf{x}|} - e^{-ik_0|\mathbf{x}|}\}\mathrm{d}|\mathbf{x}|\int_0^\pi e^{-i|\mathbf{k}||\mathbf{x}|\cos\theta}\,|\mathbf{x}|\operatorname{sen}\theta\,d\theta \\ &= 2\pi i|\mathbf{k}|^{-1}\int_0^\infty \{e^{ik_0|\mathbf{x}|} - e^{-ik_0|\mathbf{x}|}\}\;d|\mathbf{x}|\{e^{-i|\mathbf{k}||\mathbf{x}|} - e^{i|\mathbf{k}||\mathbf{x}|}\} \\ &= 2\pi i|\mathbf{k}|^{-1}\int_{-\infty}^{\infty} \{e^{i(k_0-|\mathbf{k}|)a} - e^{i(k_0+|\mathbf{k}|)a}\}\,da \\ &= 4\pi^2 i|\mathbf{k}|^{-1}\{\delta(k_0 - |\mathbf{k}|) - \delta(k_0 + |\mathbf{k}|)\} \\ &= 4\pi^2 i\Delta(k).\end{aligned} \tag{23}$$

Luego, la transformada de Fourier es la misma función con el coeficiente $4\pi^2 i$. Intercambiando en (23) k y x, resulta

$$\Delta(x) = -i/4\pi^2.\int \Delta(k)e^{ik\cdot x}\,d^4k. \tag{24}$$

Algunas de las propiedades más importantes de $\Delta(x)$ pueden deducirse de su análisis de Fourier. En primer lugar, la ecuación (24) nos permite ver que $\Delta(x)$ puede descomponerse en ondas que avanzan a la velocidad de la luz. Para obtener una ecuación que dé cuenta de este

resultado, aplicamos el operador $\Box$ a ambos miembros de (24), resultando

$$\Box\Delta(x) = -i/4\pi^2. \int \Delta(k)\Box e^{ik\cdot x}\, d^4k = i/4\pi^2. \int k_\mu k^\mu \Delta(k) e^{ik\cdot x}\, d^4k.$$

Y como $k_\mu k^\mu \Delta(k) = 0$, se tendrá

$$\Box\Delta(x) = 0. \tag{25}$$

Esta ecuación se verifica en todo el espacio-tiempo. Para dar sentido a $\Box\Delta(x)$ en los puntos en que $\Delta(x)$ es singular, hemos de tomar la integral de $\Box\Delta(x)$ extendida a un pequeño volumen cuatridimensional que rodee al punto en cuestión, y transformarla luego mediante el teorema de Gauss, en una integral de superficie tridimensional. La ecuación (25) nos dice que dicha integral de superficie tridimensional siempre es nula.

La función $\Delta(x)$ es nula sobre toda la superficie tridimensional $x_0 = 0$. Hallemos el valor de $\partial\Delta(x)/\partial x_0$, sobre dicha superficie. Evidentemente es nulo en todo punto excepto en $x_1 = x_2 = x_3 = 0$, donde presenta una singularidad que vamos a calcular a continuación. Derivemos ambos miembros de (24) respecto a x_0,

$$\begin{aligned}\partial\Delta(x)/\partial x_0 &= 1/4\pi^2 \cdot \int k_0\Delta(k) e^{ik\cdot x}\, d^4k \\ &= 1/4\pi^2 \cdot \int k_0|\mathbf{k}|^{-1}\{\delta(k_0 - |\mathbf{k}|) - \delta(k_0 + |\mathbf{k}|)\} e^{ik\cdot x}\, d^4k \\ &= 1/4\pi^2 \cdot \int \{\delta(k_0 - |\mathbf{k}|) + \delta(k_0 + |\mathbf{k}|)\} e^{ik\cdot x}\, d^4k.\end{aligned}$$

Poniendo ahora $x_0 = 0$ en ambos miembros, resulta

$$\begin{aligned}[\partial\Delta(x)/\partial x_0]_{x_0=0} &= 1/4\pi^2. \int \{\delta(k_0 - |\mathbf{k}|) + \delta(k_0 + |\mathbf{k}|)\} e^{-i(\mathbf{kx})}\, d^4k \\ &= 1/2\pi^2. \int e^{-i(\mathbf{kx})}\, d^3k \\ &= 4\pi\, \delta(x_1)\delta(x_2)\, \delta(x_3) = 4\pi\, \delta(\mathbf{x}).\end{aligned} \tag{26}$$

Luego en el punto $x_1 = x_2 = x_3 = 0$ aparece la singularidad δ ordinaria con el coeficiente 4π.

Calculemos ahora $[A_\mu(x), A_\nu(x')]$. A partir de (3), (16) y (17) tenemos

$$\begin{aligned}&[A_\mu(x), A_\nu(x')] \\ &= \int\int [A_{\mu\mathbf{k}}\, e^{ik\cdot x} + \bar{A}_{\mu\mathbf{k}}\, e^{-ik\cdot x}, A_{\nu\mathbf{k}'}\, e^{ik'\cdot x'} + \bar{A}_{\nu\mathbf{k}'}\, e^{-ik'\cdot x'}]\, d^3k d^3k' \\ &= ig_{\mu\nu}/4\pi^2 \cdot \int\int k_0^{-1}\{e^{-ik\cdot x}e^{ik'\cdot x'} - e^{ik\cdot x}e^{-ik'\cdot x'}\}\, \delta(\mathbf{k} - \mathbf{k}')\, d^3k d^3k' \\ &= ig_{\mu\nu}/4\pi^2 \cdot \int k_0^{-1}\{e^{-ik\cdot(x-x')} - e^{ik\cdot(x-x')}\}\, d^3k.\end{aligned} \tag{27}$$

El k_0 que figura aquí es por definición igual a $|\mathbf{k}|$ y, por tanto, es siempre

positivo. Poniendo $-\mathbf{k}$ en lugar de $\mathbf{k}$ en la segunda parte del integrando, deducimos que (27) es igual a la integral cuádruple

$$ig_{\mu\nu}/4\pi^2 \cdot \int |\mathbf{k}|^{-1}\{\delta(k_0 - |\mathbf{k}|) - \delta(k_0 + |\mathbf{k}|)\}e^{-ik\cdot(x-x')}\, d^4k$$
$$= ig_{\mu\nu}/4\pi^2 \cdot \int \Delta(k)e^{-ik\cdot(x-x')}\, d^4k,$$

donde k_0 puede tomar todos los valores, tanto positivos como negativos. Hallando el valor de este término con ayuda de (24), obtenemos finalmente

$$[A_\mu(x), A_\nu(x')] = g_{\mu\nu}\Delta(x - x'), \tag{28}$$

resultado que nos muestra claramente que los P.B. son invariantes frente a las transformaciones de Lorentz.

La fórmula (28) nos dice que los potenciales en dos puntos cualesquiera del espacio-tiempo conmutan siempre salvo en el caso en que la línea que une ambos puntos sea una línea nula, es decir, una trayectoria luminosa. Dicha fórmula está de acuerdo con las ecuaciones de campo $\Box A_\mu(x) = 0$, puesto que según (25), al aplicar $\Box$ al segundo miembro da cero.

76. *Variables dinámicas de Schrödinger*

Para establecer la invariancia de una teoría cuántica de campos frente a las transformaciones de Lorentz, la imagen más conveniente es la de Heisenberg, que empleábamos en las dos secciones anteriores. Pero para estudiar ejemplos concretos sigue siendo necesaria la imagen de Schrödinger. La teoría expresada en imagen de Schrödinger aparentemente no tiene forma relativista, pues considera estados en un instante de tiempo y se pregunta cómo varían dichos estados con el tiempo y, por tanto, se refiere fundamentalmente a un sistema de referencia de Lorentz particular. Sin embargo, una vez comprobado que una teoría es relativista puesta en imagen de Heisenberg, no tenemos por qué preocuparnos por la apariencia no relativista de dicha teoría puesta en imagen de Schrödinger, pues sabemos que ambas imágenes son equivalentes.

Para trabajar en imagen de Schrödinger hemos de introducir variables dinámicas adecuadas que describan el campo en un instante dado. Las cantidades $A_\mu(x)$ y $\partial A_\mu/\partial x_0$, particularizadas para todos los valores de x_1, x_2, x_3, y un valor fijo de x_0 son suficientes para determinar A_μ en todo instante con ayuda de la ecuación de campo (1); por lo tanto, dichas cantidades pueden tomarse como las variables dinámicas en un instante de

tiempo y, en consecuencia, como las variables dinámicas en la imagen de Schrödinger. Pongamos

$$B_\mu = \partial A_\mu / \partial x_0. \tag{29}$$

Entonces las variables dinámicas en la imagen de Schrödinger serán $A_{\mu\mathbf{x}}$, $B_{\mu\mathbf{x}}$, donde $\mathbf{x}$ representa x_1, x_2, x_3.

El análisis de Fourier de dichas variables es, según (3) y (9).

$$\left.\begin{aligned} A_{\mu\mathbf{x}} &= \int (A_{\mu\mathbf{k}t} + \bar{A}_{\mu-\mathbf{k}t}) e^{-i(\mathbf{k}\mathbf{x})}\, d^3k, \\ B_{\mu\mathbf{x}} &= i \int k_0 (A_{\mu\mathbf{k}t} - \bar{A}_{\mu-\mathbf{k}t}) e^{-i(\mathbf{k}\mathbf{x})}\, d^3k. \end{aligned}\right\} \tag{30}$$

Podemos invertir la transformación de Fourier y expresar $A_{\mu\mathbf{k}t} + \bar{A}_{\mu-\mathbf{k}t}$ y $A_{\mu\mathbf{k}t} - \bar{A}_{\mu-\mathbf{k}t}$ en función de $A_{\mu\mathbf{x}}$ y $B_{\mu\mathbf{x}}$ respectivamente. Luego $A_{\mu\mathbf{k}t}$ y $\bar{A}_{\mu\mathbf{k}t}$ quedan determinadas por $A_{\mu\mathbf{x}}$ y $B_{\mu\mathbf{x}}$ particularizadas para todo $\mathbf{x}$ (y un x_0 fijo), y en consecuencia, también podemos tomar como variables dinámicas en un instante de tiempo las $A_{\mu\mathbf{k}t}$, $\bar{A}_{\mu\mathbf{k}t}$, que constituyen, por tanto, un nuevo conjunto de variables dinámicas de Schrödinger.

Si queremos trabajar con las variables $A_{\mu\mathbf{x}}$, $B_{\mu\mathbf{x}}$, es necesario conocer sus P.B. Se pueden obtener bien a partir de los desarrollos de Fourier (30) y las ecuaciones (16) y (17), o bien a partir de la forma general (28) de los P.B. El segundo procedimiento nos da los resultados que queremos más directamente. Si en (28) hacemos $x'_0 = x_0$, resulta

$$[A_{\mu\mathbf{x}}, A_{\nu\mathbf{x}'}] = 0. \tag{31}$$

Derivando (28) respecto a x_0 y haciendo luego $x'_0 = x_0$, con ayuda de (26), resulta

$$[B_{\mu\mathbf{x}}, A_{\nu\mathbf{x}'}] = 4\pi g_{\mu\nu} \delta(\mathbf{x} - \mathbf{x}'). \tag{32}$$

Derivando (28) respecto a x_0 y respecto a x'_0 y haciendo luego $x'_0 = x_0$ obtenemos

$$[B_{\mu\mathbf{x}}, B_{\nu\mathbf{x}'}] = 0, \tag{33}$$

puesto que para $x_0 = 0$, $\partial^2 \Delta(x)/\partial x_0^2 = 0$. Las ecuaciones (31), (32) y (33) nos dan todos los P.B. entre las variables $A_{\mu\mathbf{x}}$, $B_{\mu\mathbf{x}}$. Según ellas vemos que, salvo coeficientes numéricos, las $A_{\mu\mathbf{x}}$ pueden ser consideradas como un conjunto de coordenadas dinámicas y las $B_{\mu\mathbf{x}}$ como momentos canónicos conjugados; en el segundo miembro de (32) aparece una función δ en lugar de un símbolo δ con dos subíndices, debido a que el número de grados de libertad es un infinito continuo.

Podemos descomponer $A_{r\mathbf{x}}$ en una parte transversal y otra longitudinal, como indican las ecuaciones (8) y (6). Para $B_{r\mathbf{x}}$ podemos hacer lo mismo, obteniéndose

$$B_r = \mathscr{B}_r + \partial U / \partial x_r \tag{34}$$

donde

$$\partial \mathscr{B}_r/\partial x_r = 0. \tag{35}$$

A partir de (7), sustituyendo **k** por $-$**k** en el segundo término del integrando, resulta

$$V = i \int k_s k_0^{-2}(A_{s\mathbf{k}t} + \bar{A}_{s-\mathbf{k}t})e^{-i(\mathbf{kx})}\, d^3k. \tag{36}$$

La ecuación correspondiente para U, dado que $U = \partial V/\partial x_0$, es

$$U = -\int k_s k_0^{-1}(A_{s\mathbf{k}t} - \bar{A}_{s-\mathbf{k}t})e^{-i(\mathbf{kx})}\, d^3k. \tag{37}$$

El campo eléctrico viene dado por

$$\mathscr{E}_r = -B_r - \frac{\partial A_0}{\partial x_r} = -\mathscr{B}_r - \frac{\partial(A_0 + U)}{\partial x_r} \tag{38}$$

Luego

$$\operatorname{div} \boldsymbol{\mathscr{E}} = -\frac{\partial B_r}{\partial x_r} - \nabla^2 A_0 = -\nabla^2(A_0 + U). \tag{39}$$

Es evidente que toda variable longitudinal conmuta con toda variable transversal. Ahora vamos a desarrollar algunos P.B. de gran utilidad. Emplearemos la notación

$$\frac{\partial f_{\mathbf{x}}}{\partial x_r} = f_{\mathbf{x}}{}^r, \quad \frac{\partial f_{\mathbf{x}'}}{\partial x'_r} f_{\mathbf{x}'}{}^{r'}. \tag{40}$$

donde $f_{\mathbf{x}}$ designa una función de campo cualquiera.

Si en (32) hacemos $\mu = r$, $\nu = s$ y derivamos la ecuación respecto a x_r, obtenemos

$$[B_{r\mathbf{x}}{}^r, A_{s\mathbf{x}'}] = 4\pi g_{rs}\delta^r(\mathbf{x} - \mathbf{x}') = -4\pi\delta^s(\mathbf{x} - \mathbf{x}'),$$

o, según (39),

$$[\operatorname{div} \boldsymbol{\mathscr{E}}_{\mathbf{x}}, A_{s\mathbf{x}'}] = 4\pi\delta^s(\mathbf{x} - \mathbf{x}'). \tag{41}$$

Ahora bien, (39) nos muestra que div $\boldsymbol{\mathscr{E}}$ es una función de las variables longitudinales únicamente, y por lo tanto, (41) da

$$[\operatorname{div} \boldsymbol{\mathscr{E}}_{\mathbf{x}}, V_{\mathbf{x}'}{}^{s'}] = 4\pi\delta^s(\mathbf{x} - \mathbf{x}') = -4\pi\delta^{s'}(\mathbf{x} - \mathbf{x}').$$

Integrando respecto a x'_s, obtenemos

$$[\operatorname{div} \boldsymbol{\mathscr{E}}_{\mathbf{x}}, V_{\mathbf{x}'}] = -4\pi\delta(\mathbf{x} - \mathbf{x}'), \tag{42}$$

donde no aparece ninguna constante de integración, ya que las funciones

de campo $\mathcal{E}_{\mathbf{x}}$ y $V_{\mathbf{x}}$ están formadas por ondas de longitud de onda no nula. De (42) y (39) resulta

$$\nabla^2[U_{\mathbf{x}}, V_{\mathbf{x}'}] = 4\pi\delta(\mathbf{x}-\mathbf{x}').$$

Integrando con ayuda de la fórmula (72) de § 38, obtenemos

$$[U_{\mathbf{x}}, V_{\mathbf{x}'}] = -|\mathbf{x}-\mathbf{x}'|^{-1}, \tag{43}$$

no apareciendo tampoco en el segundo miembro ninguna constante de integración u otros términos que no se anulen en el infinito, debido a que $U_{\mathbf{x}}$ y $V_{\mathbf{x}}$ están formadas asimismo por ondas de longitud de onda no nula. A partir de (38) y (43) tenemos

$$[\mathcal{E}_{r\mathbf{x}}, V_{\mathbf{x}'}] = -[U_{\mathbf{x}}{}^r, V_{\mathbf{x}'}] = -(x_r - x'_r)\,|\mathbf{x}-\mathbf{x}'|^{-3}. \tag{44}$$

Vamos a obtener ahora la expresión del hamiltoniano en función de las variables $A_{\mu\mathbf{x}}$ y $B_{\mu\mathbf{x}}$. De la segunda ecuación (30) resulta

$$\begin{aligned}
&\int B_{\mu\mathbf{x}} B^{\mu}{}_{\mathbf{x}}\, d^3x \\
&\quad = -\int\int\int k_0 k'_0 (A_{\mu\mathbf{k}t} - \bar{A}_{\mu-\mathbf{k}t})(A^{\mu}{}_{\mathbf{k}'t} - \bar{A}^{\mu}{}_{-\mathbf{k}'t}) e^{-i(\mathbf{k}\mathbf{x})} e^{-i(\mathbf{k}'\mathbf{x})}\, d^3k\, d^3k'\, d^3x \\
&\quad = -8\pi^3 \int\int k_0 k'_0 (A_{\mu\mathbf{k}t} - \bar{A}_{\mu-\mathbf{k}t})(A^{\mu}{}_{\mathbf{k}'t} - \bar{A}^{\mu}{}_{-\mathbf{k}'t})\delta(\mathbf{k}+\mathbf{k}')\, d^3k\, d^3k' \\
&\quad = -8\pi^3 \int k_0^2 (A_{\mu\mathbf{k}t} - \bar{A}_{\mu-\mathbf{k}t})(A^{\mu}{}_{-\mathbf{k}t} - \bar{A}^{\mu}{}_{\mathbf{k}t})\, d^3k.
\end{aligned}$$

Análogamente, de la primera ecuación (30),

$$\begin{aligned}
&\int A_{\mu\mathbf{x}}{}^r A^{\mu}{}_{\mathbf{x}}{}^r\, d^3x \\
&\quad = -\int\int\int k_r k'_r (A_{\mu\mathbf{k}t} + \bar{A}_{\mu-\mathbf{k}t})(A^{\mu}{}_{\mathbf{k}'t} + \bar{A}^{\mu}{}_{-\mathbf{k}'t}) e^{-i(\mathbf{k}\mathbf{x})} e^{-i(\mathbf{k}'\mathbf{x})}\, d^3k\, d^3k'\, d^3x \\
&\quad = 8\pi^3 \int k_0^2 (A_{\mu\mathbf{k}t} + \bar{A}_{\mu-\mathbf{k}t})(A^{\mu}{}_{-\mathbf{k}t} + \bar{A}^{\mu}{}_{\mathbf{k}t})\, d^3k.
\end{aligned}$$

Sumándolas y dividiendo por -8π, resulta

$$\begin{aligned}
-(8\pi)^{-1} \int (B_{\mu} B^{\mu} + A_{\mu}{}^r A^{\mu r})\, d^3x \\
= -2\pi^2 \int k_0^2 (A_{\mu\mathbf{k}t}\, \bar{A}^{\mu}{}_{\mathbf{k}t} + \bar{A}_{\mu-\mathbf{k}t}\, A^{\mu}{}_{-\mathbf{k}t})\, d^3k.
\end{aligned}$$

que es igual al H_F de (19) salvo un término numérico infinito. Ya vimos que la fórmula (19) para H_F dejaba en libertad un término numérico arbitrario, y así pues podemos tomar

$$H_F = -(8\pi)^{-1} \int (B_{\mu} B^{\mu} + A_{\mu}{}^r A^{\mu r})\, d^3x \tag{45}$$

con un término numérico arbitrario distinto del de (19).

Por supuesto, el hamiltoniano (45) puede utilizarse igualmente para establecer las ecuaciones del movimiento de Heisenberg, pues el término numérico arbitrario que figura en él no tiene ningún efecto. Puede comprobarse fácilmente que empleando (31), (32) y (33), resulta

$$\left.\begin{aligned} \partial A_\mu/\partial x_0 &= [A_\mu, H_F] = B_\mu, \\ \text{y} \quad \partial B_\mu/\partial x_0 &= [B_\mu, H_F] = \nabla^2 A_\mu, \end{aligned}\right\} \quad (46)$$

que coinciden con (29) y (1). Asimismo permite obtener la ecuación de movimiento de Schrödinger

$$i\hbar\, d|P\rangle/dx_0 = H_F|P\rangle$$

para un ket $|P\rangle$ que represente un estado en la imagen de Schrödinger. El término numérico arbitrario modifica a $|P\rangle$ en un factor de fase, cuestión ésta que no tiene importancia física.

Podemos descomponer la expresión (45) de H_F en una parte transversal H_{FT} y otra longitudinal H_{FL}. A partir de (34) tenemos

$$\begin{aligned} \int B_r B_r\, d^3x &= \int (\mathscr{B}_r + U^r)(\mathscr{B}_r + U^r)\, d^3x \\ &= \int \mathscr{B}_r \mathscr{B}_r\, d^3x + \int U^r U^r\, d^3x, \end{aligned}$$

ya que los términos cruzados son nulos, pues según (35),

$$\int U^r \mathscr{B}_r\, d^3x = -\int U \mathscr{B}_r{}^r\, d^3x = 0$$

Análogamente, a partir de (8), resulta

$$\int A_r{}^s A_r{}^s\, d^3x = \int \mathscr{A}_r{}^s \mathscr{A}_r{}^s\, d^3x + \int V^{rs} V^{rs}\, d^3x,$$

siendo también nulos los términos cruzados. Luego, (45) es igual a

$$H_F = H_{FT} + H_{FL},$$

con

$$H_{FT} = (8\pi)^{-1} \int (\mathscr{B}_r \mathscr{B}_r + \mathscr{A}_r{}^s \mathscr{A}_r{}^s)\, d^3x \quad (47)$$

y

$$H_{FL} = (8\pi)^{-1} \int (U^r U^r + V^{rs} V^{rs} - B_0 B_0 - A_0{}^r A_0{}^r)\, d^3x. \quad (48)$$

Tengamos en cuenta que el término

$$(8\pi)^{-1} \int \mathscr{A}_r{}^s \mathscr{A}_r{}^s\, d^3x$$

de H_{FT} puede transformarse en

$$\begin{aligned} -(8\pi)^{-1} \int \mathcal{A}_r \mathcal{A}_r{}^{ss}\, d^3x &= -(8\pi)^{-1} \int \mathcal{A}_r (\mathcal{A}_r{}^{ss} - \mathcal{A}_s{}^{rs})\, d^3x \\ &= (8\pi)^{-1} \int \mathcal{A}_r{}^s\, (\mathcal{A}_r{}^s - \mathcal{A}_s{}^r)\, d^3x \\ &= (16\pi)^{-1} \int (\mathcal{A}_r{}^s - \mathcal{A}_s{}^r)(\mathcal{A}_r{}^s - \mathcal{A}_s{}^r)\, d^3x \\ &= (8\pi)^{-1} \int \mathcal{H}^2 d^3x, \end{aligned}$$

y por lo tanto, dicho término es precisamente la energía magnética. Nuevas integraciones por partes dan

$$\int V^{rs}V^{rs}\, d^3x = \int V^{rr}V^{ss}\, d^3x,$$

y con ello (48) puede ponerse en la forma

$$H_{FL} = (8\pi)^{-1} \int \{(U - A_0)^r(U + A_0)^r + (V^{rr} - B_0)(V^{ss} + B_0)\}\, d^3x. \quad (49)$$

77. *Condiciones suplementarias*

Volvamos a la ecuación de Maxwell (2), que hasta ahora no habíamos tenido en cuenta. No podemos introducirla directamente en la teoría cuántica sin que se produzcan contradicciones. El primer miembro de la ecuación, según las condiciones cuánticas (28), no conmuta con $A_\nu(\mathbf{x}')$ y, en consecuencia, no puede ser nulo. El método de evitar la dificultad fue indicado por Fermi. Consiste en tomar una ecuación menos exigente

$$(\partial A_\mu / \partial x_\mu)|P\rangle = 0. \quad (50)$$

suponiéndola válida para todo $|P\rangle$ que corresponda a un estado que pueda darse en la naturaleza. Existe una ecuación (50) para cada punto del espacio-tiempo, y todo ket que corresponda a un estado que pueda darse en la naturaleza, habrá de verificar todas esas ecuaciones.

Las condiciones que, como (50), deben verificar los kets que corresponden a estados que pueden darse en la naturaleza las denominaremos *condiciones suplementarias*. La existencia de condiciones suplementarias en la teoría no lleva consigo ninguna desviación o modificación de los principios generales de la mecánica cuántica. El principio de superposición de los estados y toda la teoría de estados, variables dinámicas y observables dada en el capítulo II sigue siendo aplicable aunque existan condiciones suplementarias, si imponemos una nueva condición para que un operador lineal sea un observable. Diremos que un operador lineal es

físico, si tiene la propiedad de que, al aplicarlo a un ket cualquiera que verifique las condiciones suplementarias, da como resultado otro ket que también las verifica. Para que un operador lineal pueda representar un observable, además de las condiciones de § 10, tendrá que satisfacer la nueva condición de ser físico.

Ya vimos un ejemplo de condición suplementaria en la teoría de los sistemas que contienen partículas idénticas. La condición de que únicamente las funciones de onda simétricas, o únicamente las funciones de onda antisimétricas, representen estados que puedan darse en la naturaleza es precisamente una condición del tipo (50), y constituye lo que hemos llamado una condición suplementaria. En dicha teoría, la condición para que un operador lineal sea físico es que sea simétrico en las partículas.

Al introducir condiciones suplementarias en nuestra teoría, hemos de comprobar que son coherentes, es decir, que no son tan restrictivas como para que ningún ket las verifique. Si tenemos más de una condición suplementaria, podemos deducir nuevas condiciones suplementarias a partir de ellas tomando los P.B. entre los operadores que figuran en las mismas; así, si en una ocasión tenemos las condiciones

$$U|P\rangle = 0, \qquad V|P\rangle = 0, \tag{51}$$

podemos deducir

$$[U,V]|P\rangle = 0, \qquad [U,[U,V]]|P\rangle = 0, \tag{52}$$

y así sucesivamente. Para comprobar que las condiciones suplementarias son coherentes hemos de considerar todas las que se pueden obtener por este procedimiento y ver que pueden ser satisfechas a la vez; lo que corrientemente se consigue demostrando que tras un cierto número de procesos de este tipo, las nuevas condiciones suplementarias se verifican idénticamente o bien son repeticiones de las ya obtenidas.

Asimismo hemos de comprobar que las condiciones suplementarias están de acuerdo con las ecuaciones del movimiento. En la imagen de Heisenberg, en la que el ket $|P\rangle$ de (51) es fijo, tendremos distintas condiciones suplementarias relativas a los distintos instantes, y todas ellas tienen que ser coherentes en el sentido que hemos indicado. En la imagen de Schrödinger, el ket $|P\rangle$ varía con el tiempo según la ecuación de Schrödinger, y hemos de exigir que si $|P\rangle$ verifica las condiciones suplementarias en el instante inicial las siga verificando también en todo instante. Esto es equivalente a exigir que $d|P\rangle/dt$ satisfaga las condiciones suplementarias, o sea, que $H|P\rangle$ las satisfaga, que es lo mismo que exigir que H sea físico.

Cuando se tiene una condición suplementaria $U|P\rangle = 0$, es conveniente escribir

$$U \approx 0 \tag{53}$$

diciendo que (53) es una ecuación débil, para distinguirla de una ecuación ordinaria o fuerte. Multiplicando una ecuación débil a la izquierda por un factor obtenemos una nueva ecuación débil, pero, en general, si el factor se multiplica por la derecha no se obtiene una ecuación consecuencia de la anterior. Luego, las ecuaciones débiles no deben emplearse para calcular los P.B. Con esta nomenclatura, la condición (52) para que las condiciones suplementarias sean coherentes equivale a la condición de que los P.B. de los operadores que determinan las condiciones suplementarias sean nulos débilmente.

La condición para que una variable dinámica ξ sea física es que para cada ecuación suplementaria $U|P\rangle = 0$ se tenga

$$U\xi|P\rangle = 0,$$

y por tanto, que

$$[U, \xi]|P\rangle = 0.$$

Luego, la condición es que el P.B. de la variable dinámica con cada uno de los operadores que determinan las condiciones suplementarias sea nulo débilmente.

Volvamos a la electrodinámica. Consideramos la ecuación (2) como débil; por tanto, debemos escribirla

$$\partial A_\mu/\partial x_\mu \approx 0. \tag{54}$$

En la imagen de Heisenberg tenemos una de estas ecuaciones para cada punto x. Para ver que son coherentes, tomamos dos puntos x y x' arbitrarios del espacio-tiempo y construimos el P.B.

$$\left[\frac{\partial A_\mu(x)}{\partial x_\mu}, \frac{\partial A_\nu(x')}{\partial x'_\nu}\right] = \frac{\partial^2}{\partial x_\mu \partial x'_\nu}[A_\mu(x), A_\nu(x')].$$

Calculándolo con la ayuda de (28), resulta

$$g_{\mu\nu}\frac{\partial^2\Delta(x-x')}{\partial x_\mu \partial x'_\nu} = -\Box\Delta(x-x') = 0$$

según (25); luego, las condiciones de coherencia se verifican fuertemente. Al comprobar que las condiciones suplementarias son coherentes en todo instante en la imagen de Heisenberg, queda comprobado que están de acuerdo con las ecuaciones de movimiento.

Puesto que (54) es una ecuación débil, todas sus consecuencias de la teoría de Maxwell ordinaria serán válidas en la teoría cuántica únicamente como ecuaciones débiles. Las ecuaciones

$$\operatorname{div} \mathcal{H} = 0, \qquad \partial\mathcal{H}/\partial t = -\operatorname{rot} \mathcal{E}$$

son consecuencia inmediata de las definiciones de $\mathcal{E}$ y $\mathcal{H}$ en función de

los potenciales y, por lo tanto, serán válidas fuertemente también en la teoría cuántica. Las otras ecuaciones de Maxwell en el vacío

$$\operatorname{div} \mathcal{E} \approx 0, \qquad \partial\mathcal{E}/\partial t \approx \operatorname{rot} \mathcal{H}, \tag{55}$$

en teoría cuántica serán ecuaciones débiles, pues para obtenerlas es necesario emplear tanto (54) como (1).

Las cantidades de campo $\mathcal{E}$ y $\mathcal{H}$ son componentes del tensor antisimétrico $\partial A^\nu/\partial x_\mu - \partial A^\mu/\partial x_\nu$. El P.B. de dicho tensor con el operador de (54) en un punto x' general es

$$\left[\frac{\partial A^\nu(x)}{\partial x_\mu} - \frac{\partial A^\mu(x)}{\partial x_\nu}, \frac{\partial A_\sigma(x')}{\partial x'_\sigma}\right] = g^\nu_\sigma \frac{\partial^2\Delta(x-x')}{\partial x_\mu \partial x'_\sigma} - g^\mu_\sigma \frac{\partial^2\Delta(x-x')}{\partial x_\nu \partial x'_\sigma} = 0.$$

Por consiguiente, $\mathcal{E}$ y $\mathcal{H}$ son físicos. Los potenciales A_μ no son físicos.

Las condiciones suplementarias que pesan sobre las variables dinámicas en un instante particular son

$$\frac{\partial A_\mu}{\partial x_\mu} \approx 0, \qquad \frac{\partial}{\partial x_0}\frac{\partial A_\mu}{\partial x_\mu} \approx 0. \tag{56}$$

Nuevas derivaciones respecto a x_0 no dan lugar a ecuaciones independientes, sino a ecuaciones que son consecuencia de éstas y de la ecuación fuerte (1). Las condiciones suplementarias, puestas en función de las variables de Schrödinger de § 76, son

$$B_0 + A_r{}^r \approx 0 \tag{57}$$

y

$$(A_0{}^r + B_r)^r \approx 0. \tag{58}$$

La ecuación (58) es igual a la primera ecuación (55) y, según (39), puede escribirse también

$$\nabla^2(A_0 + U) \approx 0$$

Puesto que esta ecuación se verifica sobre todo el espacio tridimensional, resulta

$$A_0 + U \approx 0. \tag{59}$$

Llamando $A_r{}^r = V^{rr}$, de (49) resulta

$$H_{FL} \approx 0. \tag{60}$$

Luego, *para los estados que se dan en la naturaleza la energía del campo longitudinal es nula.*

Con el fin de establecer una representación conveniente, introducimos un ket standard $|0_F\rangle$ que verifica las condiciones suplementarias

$$(B_0 + A_r{}^r)|0_F\rangle = 0, \qquad (A_0 + U)|0_F\rangle = 0, \tag{61}$$

y además

$$\bar{\mathscr{A}}_{r\mathbf{k}}|0_F\rangle = 0. \tag{62}$$

Estas ecuaciones son coherentes, pues $\bar{\mathscr{A}}_{r\mathbf{k}}$ conmuta con los operadores de (61), y además son suficientes para determinar $|0_F\rangle$ completamente, salvo un factor numérico, ya que las únicas variables dinámicas independientes que tenemos son A_0, B_0, U, A_r^r, $\mathscr{A}_{r\mathbf{k}}$, $\bar{\mathscr{A}}_{r\mathbf{k}}$ y de ellas $A_0 + U$, $B_0 + A_r^r$, $\bar{\mathscr{A}}_{r\mathbf{k}}$ constituyen un conjunto completo. Con ayuda de este ket standard podemos expresar cualquier ket en la forma

$$\Psi(A_0, B_0, \mathscr{A}_{r\mathbf{k}})|0_F\rangle. \tag{63}$$

Esta representación es exactamente la representación de Fock en lo que concierne a las variables dinámicas transversales $\mathscr{A}_{r\mathbf{k}}$, $\bar{\mathscr{A}}_{r\mathbf{k}}$ y, por tanto, Ψ ha de ser de la forma de una serie de potencias en las variables $\mathscr{A}_{r\mathbf{k}}$, cuyos términos corresponden a los distintos números de fotones presentes. El número de variables que figuran en la expresión de Ψ es un infinito de la potencia del continuo, luego, Ψ es lo que en matemáticas se denomina un 'funcional'.

Si el ket (63) verifica las condiciones suplementarias, Ψ ha de ser independiente de A_0 y B_0, y luego, ha de ser función únicamente de $\mathscr{A}_{r\mathbf{k}}$. Por tanto, los estados físicos están representados por kets de la forma

$$\Psi(\mathscr{A}_{r\mathbf{k}})|0_F\rangle, \tag{64}$$

siendo Ψ una serie de potencias en las variables $\mathscr{A}_{r\mathbf{k}}$. El ket standard $|0_F\rangle$ representa el estado físico que corresponde a la no existencia de fotones, es decir, el vacío absoluto.

Nuestro hamiltoniano H_F y sus partes H_{FL}, H_{FT} contenían hasta ahora términos numéricos arbitrarios. Es conveniente elegir dichos términos de modo que, en el vacío absoluto, tanto H_{FL} como H_{FT} sean nulos. La conclusión (60) indica que en el H_{FL} dado por (48) o (49) ya hemos elegido correctamente el término numérico para que H_{FL} valga cero en el estado de vacío absoluto y así como en los demás estados físicos. Si queremos que no exista punto cero de energía para los fotones, hemos de elegir el término numérico de la parte transversal de (19) de modo que H_{FT} valga

$$H_{FT} = 4\pi^2 \int k_0^2 \mathscr{A}_{r\mathbf{k}t} \bar{\mathscr{A}}_{r\mathbf{k}t} d^3k, \tag{65}$$

(47) difiere de (65) en un término numérico infinito que consiste en medio cuanto de energía para cada estado fotónico.

78. *Electrones y positrones aislados*

Vamos a considerar ahora electrones y positrones en ausencia de campo electromagnético. El estado de un electrón viene dado, como en el capí-

tulo XI, por una función de onda ψ con cuatro componentes ψ ($a = 1$, 2, 3, 4), que verifica la ecuación de ondas

$$i\hbar \frac{\partial \psi}{\partial x_0} = -i\hbar\, \alpha_r \frac{\partial \psi}{\partial x_r} + \alpha_m\, m\psi. \tag{66}$$

Para obtener una teoría aplicable a un conjunto de electrones, aplicaremos el método de segunda cuantificación de § 65, que consistía en sustituir la función de onda de un electrón por un conjunto de operadores que verifican ciertas relaciones de anticonmutación.

Cuando consideremos la ψ en distintos puntos en un instante dado escribiremos $\psi_{\mathbf{x}}$ donde $\mathbf{x}$ representa x_1, x_2, x_3. Sus componentes serán $\psi_{a\mathbf{x}}$. Mediante una transformación de Fourier pasamos a la representación de momentos cuya función de onda es $\psi_{\mathbf{p}}$. Se tiene

$$\psi_{\mathbf{x}} = h^{-\frac{3}{2}} \int e^{i(\mathbf{xp})/\hbar} \psi_{\mathbf{p}}\, d^3p, \qquad \psi_{\mathbf{p}} = h^{-\frac{3}{2}} \int e^{-i(\mathbf{xp})/\hbar} \psi_{\mathbf{x}}\, d^3x. \tag{67}$$

$\psi_{\mathbf{p}}$ tendrá también cuatro componentes $\psi_{a\mathbf{p}}$, en correspondencia con las cuatro componentes de $\psi_{\mathbf{x}}$. En esta representación el operador energía es

$$p_0 = \alpha_r\, p_r + \alpha_m\, m,$$

en el que los operadores momento p_r son factores multiplicativos.

Podemos separar ψ en una parte ξ de energía positiva y otra ζ de energía negativa.

$$\psi = \xi + \zeta,$$

donde ξ y ζ tendrán cada una cuatro componentes igual que ψ. En la representación de momentos vendrán dadas por

$$\xi_{\mathbf{p}} = \frac{1}{2}\left\{1 + \frac{\alpha_r\, p_r + \alpha_m\, m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}\psi_{\mathbf{p}}, \qquad \zeta_{\mathbf{p}} = \frac{1}{2}\left\{1 - \frac{\alpha_r\, p_r + \alpha_m\, m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}\psi_{\mathbf{p}}, \tag{68}$$

ya que dichas ecuaciones nos llevan a

$$\begin{aligned} p_0\, \xi_{\mathbf{p}} &= (\alpha_r\, p_r + \alpha_m\, m)\xi_{\mathbf{p}} = \tfrac{1}{2}\{\alpha_r\, p_r + \alpha_m\, m + (\mathbf{p}^2 + m^2)^{\frac{1}{2}}\}\psi_{\mathbf{p}} \\ &= (\mathbf{p}^2 + m^2)^{\frac{1}{2}}\xi_{\mathbf{p}}, \end{aligned}$$

y análogamente a

$$p_0\, \zeta_{\mathbf{p}} = -(\mathbf{p}^2 + m^2)^{\frac{1}{2}}\zeta_{\mathbf{p}},$$

que nos dicen que $\xi_{\mathbf{p}}$ y $\zeta_{\mathbf{p}}$ son autofunciones de p_0 cuyos autovalores son respectivamente $(\mathbf{p}^2 + m^2)^{\frac{1}{2}}$ y $-(\mathbf{p}^2 + m^2)^{\frac{1}{2}}$. Trabajando con los operadores

$$\frac{1}{2}\left\{1 + \frac{\alpha_r\, p_r + \alpha_m\, m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}, \qquad \frac{1}{2}\left\{1 - \frac{\alpha_r\, p_r + \alpha_m\, m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\},$$

hemos de tener en cuenta que sus cuadrados son iguales a sí mismos, y que su producto en cualquier orden es igual a cero.

La segunda cuantificación transforma las ψ en operadores del tipo de las $\bar{\eta}$ de § 65, que verifican unas relaciones de anticonmutación del tipo de (11′) de § 65. Empleando la siguiente notación del anticonmutador

$$MN + NM = [M, N]_+, \tag{69}$$

resulta

$$\left.\begin{aligned} [\psi_{a\mathbf{x}}, \psi_{b\mathbf{x}'}]_+ &= 0, \qquad [\bar{\psi}_{a\mathbf{x}}, \bar{\psi}_{b\mathbf{x}'}]_+ = 0, \\ [\psi_{a\mathbf{x}}, \bar{\psi}_{b\mathbf{x}'}]_+ &= \delta_{ab}\,\delta(\mathbf{x} - \mathbf{x}'), \end{aligned}\right\} \tag{70}$$

apareciendo en la última ecuación la función $\delta(\mathbf{x} - \mathbf{x}')$ debido a que las $\mathbf{x}$ toman valores en un dominio continuo. Pasando a la representación $\mathbf{p}$, mediante (67), obtenemos

$$\left.\begin{aligned} [\psi_{a\mathbf{p}}, \psi_{b\mathbf{p}'}]_+ &= 0, \qquad [\bar{\psi}_{a\mathbf{p}}, \bar{\psi}_{b\mathbf{p}'}]_+ = 0, \\ [\psi_{a\mathbf{p}}, \bar{\psi}_{b\mathbf{p}'}]_+ &= \delta_{ab}\,\delta(\mathbf{p} - \mathbf{p}'). \end{aligned}\right\} \tag{71}$$

Definiendo también ahora ξ y ζ por (68), la última ecuación (71) da

$$\begin{aligned} [\xi_{a\mathbf{p}}, \bar{\xi}_{b\mathbf{p}'}]_+ &= \frac{1}{2}\left\{1 + \frac{\alpha_r p_r + \alpha_m m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}_{ac} [\psi_{c\mathbf{p}}, \bar{\psi}_{d\mathbf{p}'}]_+ \frac{1}{2}\left\{1 + \frac{\alpha_s p'_s + \alpha_m m}{(\mathbf{p}'^2 + m^2)^{\frac{1}{2}}}\right\}_{db} \\ &= \frac{1}{2}\left\{1 + \frac{\alpha_r p_r + \alpha_m m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}_{ab} \delta(\mathbf{p} - \mathbf{p}') \end{aligned} \tag{72}$$

y análogamente

$$[\zeta_{a\mathbf{p}}, \bar{\zeta}_{b\mathbf{p}'}]_+ = \frac{1}{2}\left\{1 - \frac{\alpha_r p_r + \alpha_m m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}_{ab} \delta(\mathbf{p} - \mathbf{p}') \tag{73}$$

y

$$[\xi_{a\mathbf{p}}, \bar{\zeta}_{b\mathbf{p}'}]_+ = [\bar{\xi}_{a\mathbf{p}}, \zeta_{b\mathbf{p}'}]_+ = 0.$$

Según la interpretación de § 65, los operadores $\psi_{a\mathbf{p}}$ son operadores de aniquilación de un electrón de momento $\mathbf{p}$ y los $\bar{\psi}_{a\mathbf{p}}$ son operadores de creación de un electrón de momento $\mathbf{p}$. A fin de evitar la noción no física de electrones de energía negativa, hemos de pasar a una nueva interpretación basada en la teoría de los positrones de § 73. La aniquilación de un electrón de energía negativa debe ser considerada como la creación de un hueco en el mar de electrones de energía negativa, es decir, como la creación de un positrón. Con ello, los operadores $\zeta_{a\mathbf{p}}$ se convierten en operadores de creación de un positrón. El positrón tiene un momento $-\mathbf{p}$, pues aniquila un momento $\mathbf{p}$. Análogamente los $\bar{\zeta}_{a\mathbf{p}}$ se convierten en operadores de aniquilación de un positrón de momento $-\mathbf{p}$. Los $\xi_{a\mathbf{p}}$ y $\bar{\xi}_{a\mathbf{p}}$ son respectivamente operadores de aniquilación y creación de un electrón ordinario de energía positiva y momento $\mathbf{p}$.

Observemos que aunque $\xi_{\mathbf{p}}$ tenga cuatro componentes, únicamente dos de ellas son independientes, ya que las cuatro están relacionadas por

$$\left\{1-\frac{\alpha_r p_r+\alpha_m m}{(\mathbf{p}^2+m^2)^{\frac{1}{2}}}\right\}\xi_{\mathbf{p}}=0,$$

que implica dos ecuaciones independientes. Los dos componentes de $\xi_{\mathbf{p}}$ independientes corresponden a la aniquilación de un electrón en cada uno de los dos estados de spin independientes. Análogamente $\zeta_{\mathbf{p}}$ tiene únicamente dos componentes independientes a causa de las ecuaciones

$$\left\{1+\frac{\alpha_r p_r+\alpha_m m}{(\mathbf{p}^2+m^2)^{\frac{1}{2}}}\right\}\zeta_{\mathbf{p}}=0,$$

y dichas componentes corresponden a la creación de un positrón en cada uno de los dos estados de spin independientes.

El estado vacío en el que no existen ni electrones ni positrones está representado por el ket $|0_{\mathbf{p}}\rangle$ que verifica

$$\xi_{a\mathbf{p}}|0_P\rangle=0, \qquad \zeta_{a\mathbf{p}}|0_P\rangle=0. \tag{74}$$

Podemos emplear dicho ket como ket standard de una representación. En ese caso cualquier ket se expresará en la forma

$$\Psi(\xi_{a\mathbf{p}},\zeta_{a\mathbf{p}})|0_P\rangle,$$

donde la función, o mejor dicho el funcional Ψ, es una serie de potencias en las variables $\xi_{a\mathbf{p}}$, $\zeta_{a\mathbf{p}}$. Cada término de Ψ es análogo a (17′) de § 65. No puede contener ninguna de sus variables elevada a una potencia superior a uno. Corresponde a un estado en el que existen electrones (de energía positiva) y positrones, en estados determinados por los símbolos de las variables que en ella figuran.

Según (12′) de § 65, el número total de electrones es igual a la suma de $\int \bar{\psi}_{a\mathbf{p}}\psi_{a\mathbf{p}}\,d^3p$ para todo a. Escrito en la notación de la ecuación (12) de § 67 es $\int \bar{\psi}^\dagger_{\mathbf{p}}\psi_{\mathbf{p}}d^3p$. Pasando a la representación $\mathbf{x}$ mediante (67), obtenemos

$$h^{-3}\int\int\int e^{i(\mathbf{xp})/\hbar}\,e^{-i(\mathbf{x'p'})/\hbar}\bar{\psi}^\dagger_{\mathbf{x}}\,\psi_{\mathbf{x'}}\,d^3xd^3x'd^3p=\int\bar{\psi}^\dagger_{\mathbf{x}}\,\psi_{\mathbf{x}}\,d^3x,$$

que nos dice que la densidad de electrones es $\bar{\psi}^\dagger_{\mathbf{x}}\psi_{\mathbf{x}}$. Este resultado incluye una constante infinita que representa el mar de electrones de energía negativa.

Obtenemos una cantidad con mayor significado físico tomando la carga

total Q, igual al número de electrones de energía positiva menos el número de huecos o positrones multiplicado por $-e$. Es decir,

$$Q = -e \int (\xi_{\mathbf{p}}^{\dagger}\xi_{\mathbf{p}} - \zeta_{\mathbf{p}}^{\dagger}\zeta_{\mathbf{p}})\, d^3p. \tag{75}$$

Podemos calcular esta expresión con ayuda de (68). Empleando la ecuación transpuesta de la segunda ecuación (68), es decir,

$$\zeta_{\mathbf{p}}^{\dagger} = \tfrac{1}{2}\psi_{\mathbf{p}}^{\dagger}\left\{1 - \frac{\alpha_r^{\dagger} p_r + \alpha_m^{\dagger} m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}$$

resulta

$$Q = -e\int\left\{\bar{\psi}_{\mathbf{p}}^{\dagger}\frac{1}{2}\left(1 + \frac{\alpha_r p_r + \alpha_m m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right)\psi_{\mathbf{p}} - \psi_{\mathbf{p}}^{\dagger}\frac{1}{2}\left(1 - \frac{\alpha_r^{\dagger} p_r + \alpha_m^{\dagger} m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right)\bar{\psi}_{\mathbf{p}}\right\} d^3p.$$

Por otro lado, para toda matriz α cuya suma de elementos diagonales sea igual a cero, las relaciones de anticonmutación (71) dan

$$\psi_{\mathbf{p}}^{\dagger}\alpha\psi_{\mathbf{p}'} + \psi_{\mathbf{p}}^{\dagger}\alpha^{\dagger}\bar{\psi}_{\mathbf{p}} = \alpha_{ab}(\bar{\psi}_{a\mathbf{p}}\psi_{b\mathbf{p}'} + \psi_{b\mathbf{p}'}\bar{\psi}_{a\mathbf{p}} = \alpha_{aa}\,\delta(\mathbf{p} - \mathbf{p}') = 0, \tag{76}$$

resultado que podemos suponer cierto también para $\mathbf{p}' = \mathbf{p}$. Con ello la expresión de Q se reduce a

$$Q = -e \int \tfrac{1}{2}(\bar{\psi}_{\mathbf{p}}^{\dagger}\psi_{\mathbf{p}} - \psi_{\mathbf{p}}^{\dagger}\bar{\psi}_{\mathbf{p}})\, d^3p.$$

Pasándola como antes a la representación $\mathbf{x}$, resulta

$$Q = -e \int \tfrac{1}{2}(\bar{\psi}_{\mathbf{x}}^{\dagger}\psi_{\mathbf{x}} - \psi_{\mathbf{x}}^{\dagger}\bar{\psi}_{\mathbf{x}})d^3x,$$

que nos dice que la densidad de carga es

$$j_{0\mathbf{x}} = -\tfrac{1}{2}e(\bar{\psi}_{\mathbf{x}}^{\dagger}\psi_{\mathbf{x}} - \psi_{\mathbf{x}}^{\dagger}\bar{\psi}_{\mathbf{x}}). \tag{77}$$

Además de la densidad de probabilidad $\bar{\psi}^{\dagger}\psi$, la interpretación de la función de onda de un electrón dada en § 68 determina también una corriente de probabilidad $\bar{\psi}^{\dagger}\alpha_r\psi$. En correspondencia con ello, para la segunda cuantificación tendremos un flujo de electrones determinado por el operador $\bar{\psi}_{\mathbf{x}}^{\dagger}\alpha_r\psi_{\mathbf{x}}$. El mar de electrones de energía negativa, por simetría, no da lugar a ningún flujo de electrones resultante y, por consiguiente, la corriente eléctrica vale

$$j_{r\mathbf{x}} = -e\bar{\psi}_{\mathbf{x}}^{\dagger}\alpha_r\psi_{\mathbf{x}}. \tag{78}$$

Según la fórmula (29) de § 60, válida también para fermiones, la energía total de los electrones es

$$H_{P'} = \int \bar{\psi}_{\mathbf{p}}^{\dagger} p_0 \psi_{\mathbf{p}} \, d^3p = \int \bar{\psi}_{\mathbf{p}}^{\dagger} (\alpha_r p_r + \alpha_m m) \psi_{\mathbf{p}} \, d^3p. \tag{79}$$

Puesta en representación $\mathbf{x}$ es

$$H_{P'} = \int \bar{\psi}_{\mathbf{x}}^{\dagger} (-i\hbar \alpha_r \psi_{\mathbf{x}}{}^r + \alpha_m m \psi_{\mathbf{x}}) \, d^3x. \tag{80}$$

Esta expresión de la energía total contiene un término numérico infinito que representa el mar de electrones de energía negativa.

Obtenemos una cantidad de mayor significado físico considerando la energía de todos los electrones y de todos los positrones, y tomando la energía del vacío igual a cero. Dicha cantidad es

$$H_P = \int (\mathbf{p}^2 + m^2)^{\frac{1}{2}} (\xi_{\mathbf{p}}^{\dagger} \xi_{\mathbf{p}} + \zeta_{\mathbf{p}}^{\dagger} \zeta_{\mathbf{p}}) \, d^3p \tag{81}$$

$$= \int (\mathbf{p}^2 + m^2)^{\frac{1}{2}} \left\{ \bar{\psi}_{\mathbf{p}}^{\dagger} \frac{1}{2} \left(1 + \frac{\alpha_r p_r + \alpha_m m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}} \right) \psi_{\mathbf{p}} + \right.$$

$$\left. + \psi_{\mathbf{p}}^{\dagger} \frac{1}{2} \left(1 - \frac{\alpha_r^{\dagger} p_r + \alpha_m^{\dagger} m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}} \right) \bar{\psi}_{\mathbf{p}} \right\} d^3p$$

$$= \int \tfrac{1}{2} \{ \bar{\psi}_{\mathbf{x}}^{\dagger} (\alpha_r p_r + \alpha_m m) \psi_{\mathbf{p}} - \psi_{\mathbf{p}}^{\dagger} (\alpha_r^{\dagger} p_r + \alpha_m^{\dagger} m) \bar{\psi}_{\mathbf{p}} \} \, d^3p +$$

$$+ \int (\mathbf{p}^2 + m^2)^{\frac{1}{2}} \tfrac{1}{2} (\bar{\psi}_{\mathbf{p}}^{\dagger} \psi_{\mathbf{p}} + \psi_{\mathbf{p}}^{\dagger} \bar{\psi}_{\mathbf{p}}) \, d^3p. \tag{82}$$

Según (76), la primera integral de (82) es idéntica a (79) y por consiguiente, es igual a $H_{P'}$. La segunda es una constante infinita, que es igual a menos la energía de todos los electrones de energía negativa de la distribución de vacío.

Podemos tomar como hamiltoniano bien H_P o bien $H_{P'}$. Luego, la ecuación de movimiento de Heisenberg para $\psi_{a\mathbf{x}}$ es

$$\partial \psi_{a\mathbf{x}} / \partial x_0 = [\psi_{a\mathbf{x}}, H_P] = [\psi_{a\mathbf{x}}, H_{P'}],$$

y si a partir de ella seguimos el desarrollo, volvemos a obtener la ecuación de onda (66) para ψ.

Hemos de considerar ahora la cuestión de si la teoría es o no relativista. Está dada en función de operadores ψ que verifican las ecuaciones de campo (66). Cada una de dichas ecuaciones es de la forma de la ecuación de onda de un electrón, y como ya vimos, si las ψ se transforman según la ley (20) del capítulo XI, son invariantes frente a las transformaciones de Lorentz. A este respecto, la teoría presente va más allá que la teoría de un

solo electrón, por el hecho de haber introducido relaciones de anticonmutación para las ψ y $\bar{\psi}$, y es necesario comprobar que dichas relaciones son invariantes frente a las transformaciones de Lorentz.

Vamos a proceder análogamente a como hicimos en § 75. Consideremos dos puntos cualesquiera x y x' del espacio-tiempo y formemos el anticonmutador

$$K_{ab}(x, x') = \psi_a(x)\bar{\psi}_b(x') + \bar{\psi}_b(x')\psi_a(x). \tag{83}$$

Podemos calcularlo directamente a partir de las relaciones de anticonmutación (71) de las componentes de Fourier de ψ y $\bar{\psi}$. Otro método más sencillo es tener en cuenta ciertas propiedades que ha de tener $K_{ab}(x, x')$, que vamos a enumerar a continuación:

(i) solamente depende de x_μ y x'_μ a través de su diferencia $x_\mu - x'$.

(ii) verifica la ecuación de onda

$$\left(i\hbar\frac{\partial}{\partial x_0} + i\hbar\alpha_r\frac{\partial}{\partial x_r} - \alpha_m m\right)_{ab} K_{bc}(x, x') = 0 \tag{84}$$

que se deduce teniendo en cuenta que $\psi(x)$ verifica (66);

(iii) para $x_0 = x'_0$ tiene el valor $\delta_{ab}\delta(\mathbf{x} - \mathbf{x}')$, como puede verse a partir de la tercera ecuación (70).

Estas ecuaciones son suficientes para determinar completamente $K_{ab}(x,x')$, pues (iii) determina su valor para $x_0 = x'_0$, (ii) nos da su dependencia respecto a x_0, y a partir de ello (i) nos da la dependencia respecto a x'_0. Puede verse fácilmente que la solución es

$$K_{ab}(x, x') = h^{-3} \int \sum \tfrac{1}{2}\{1 + (\alpha_r p_r + \alpha_m m)/p_0\}_{ab}\, e^{-i(x-x')\cdot p/\hbar}\, d^3p, \tag{85}$$

donde la $\sum$ se refiere a los valores $\pm(\mathbf{p}^2 + m^2)^{\frac{1}{2}}$ de p_0 para valores particulares de p_1, p_2, p_3. Verifica la condición (ii), pues al aplicarle el operador de (84) aparece en el integrando de (85) el factor $(p_0 - \alpha_r p_r - \alpha_m m)$, que multiplicado a la izquierda por el factor entre paréntesis $\{\}$ da cero. También verifica (iii), pues para $x_0 = x'_0$, la suma respecto a p_0 elimina el segundo término que figura dentro del paréntesis $\{\}$.

La ley de transformación de ψ y $\bar{\psi}$ dada en § 68 hace que las cantidades $\bar{\psi}^\dagger(x')\alpha_\mu\psi(x)$ se transformen como las cuatro componentes de un cuadrivector, y que $\bar{\psi}^\dagger(x')\alpha_m\psi(x)$ sea invariante. Luego para cualquier cuadrivector l^μ y cualquier escalar S,

$$l^\mu\bar{\psi}^\dagger(x')\alpha_\mu\psi(x) + S\bar{\psi}^\dagger(x')\alpha_m\psi(x) \tag{86}$$

es invariante. La invariancia de (86) tiene que ser una condición suficiente para asegurar la ley de transformación correcta para ψ y $\bar{\psi}$, pues nos permite deducir la invariancia de la ecuación de onda en ψ tomando $l^\mu = i\hbar\partial/\partial x_\mu$, $S = -m$.

La invariancia de (86) implica la invariancia de

$$(l^\mu \alpha_\mu + S\alpha_m)_{ab}\{\bar{\psi}_a(x')\psi_b(x) + \psi_b(x)\bar{\psi}_a(x')\}.$$

Luego,

$$(l^\mu \alpha_\mu + S\alpha_m)_{ab} K_{ba}(x, x') \tag{87}$$

ha de ser invariante para $K_{ab}(x, x')$ dado por (85), y su invariancia sería una condición suficiente para asegurar la invariancia de las relaciones de anticonmutación. Para (87) resulta

$$\begin{aligned} & h^{-3} \int \sum \tfrac{1}{2}(l^\mu \alpha_\mu + S\alpha_m)_{ab}(p_0 + \alpha_r p_r + \alpha_m m)_{ba}\, e^{-i(x-x')\cdot p/\hbar} p_0^{-1}\, d^3p \\ & = h^{-3} \int \sum \tfrac{1}{2}\{(l_0 - l_s\alpha_s + S\alpha_m)(p_0 + \alpha_r p_r + \alpha_m m)\}_{aa}\, e^{-i(x-x')\cdot p/\hbar} p_0^{-1}\, d^3p \\ & = h^{-3} \int \sum 2(l_0 p_0 - l_r p_r + Sm) e^{-i(x-x')\cdot p/\hbar} p_0^{-1}\, d^3p. \end{aligned} \tag{88}$$

Esta expresión es invariante frente a las transformaciones de Lorentz, debido a la invariancia de $p_0^{-1} d^3p$. Por tanto, queda comprobada la invariancia relativista de la teoría.

79. *La interacción*

El hamiltoniano completo para electrones y positrones en interacción con el campo electromagnético es

$$H = H_F + H_P + H_Q, \tag{89}$$

donde H_F es el hamiltoniano del campo electromagnético aislado, que viene dado por (19) o (45), H_P es el hamiltoniano de los electrones y positrones aislados, dado por (80) u (81), y H_Q es la energía de interacción, que depende tanto de las variables dinámicas de los electrones y positrones como de las del campo electromagnético. Tomamos

$$H_Q = \int A^\mu j_\mu\, d^3x, \tag{90}$$

donde j_μ viene dado por (77) y (78) y, como veremos, con esta elección se obtienen las ecuaciones de movimiento correctas. Así pues, salvo términos numéricos infinitos,

$$\begin{aligned} H = \int \{\bar{\psi}^\dagger \alpha_r(-i\hbar\psi^r - eA^r\psi) + \bar{\psi}^\dagger \alpha_m m\psi - \tfrac{1}{2}eA^0(\bar{\psi}^\dagger\psi - \psi^\dagger\bar{\psi})\}\, d^3x - \\ - (8\pi)^{-1} \int (B_\mu B^\mu + A_\mu{}^r A^{\mu r})\, d^3x. \end{aligned} \tag{91}$$

Veamos cuáles son las ecuaciones de movimiento de Heisenberg que se obtienen con el hamiltoniano (91). Tenemos

$$\begin{aligned} i\hbar\, \partial\psi_{a\mathbf{x}}/\partial x_0 &= \psi_{a\mathbf{x}} H - H\psi_{a\mathbf{x}} = \psi_{a\mathbf{x}}(H_P + H_Q) - (H_P + H_Q)\psi_{a\mathbf{x}} \\ &= \int [\psi_{a\mathbf{x}}, \bar{\psi}_{b\mathbf{x}'}]_+ \{\alpha_r(-i\hbar\psi_{\mathbf{x}'}{}^{r'} - eA^r{}_{\mathbf{x}'}\psi_{\mathbf{x}'}) + \\ &\qquad + \alpha_m m\psi_{\mathbf{x}'} - eA^0{}_{\mathbf{x}'}\psi_{\mathbf{x}'}\}_b\, d^3x' \\ &= \{\alpha_r(-i\hbar\psi_{\mathbf{x}}{}^r - eA^r{}_{\mathbf{x}}\psi_{\mathbf{x}}) + \alpha_m m\psi_{\mathbf{x}} - eA^0{}_{\mathbf{x}}\psi_{\mathbf{x}}\}_a. \end{aligned}$$

Luego,

$$\left\{\alpha_\mu\left(i\hbar\frac{\partial}{\partial x_\mu} + eA^\mu\right) - \alpha_m m\right\}\psi = 0. \tag{92}$$

que está de acuerdo con la ecuación de onda (11) del capítulo XI para un electrón. Como H es real, la ecuación de movimiento para $\bar{\psi}$ será la conjugada de la ecuación de movimiento para ψ y, por tanto, coincidirá con (12) del capítulo XI. Por consiguiente, la interacción (90) da correctamente la acción del campo sobre los electrones y positrones. Por otro lado, haciendo uso de los P.B. (46)

$$\begin{aligned} \partial A_\mu/\partial x_0 &= [A_\mu, H] = [A_\mu, H_F] \\ &= B_\mu \end{aligned} \tag{93}$$

y

$$\begin{aligned} \partial B_{\mu\mathbf{x}}/\partial x_0 &= [B_{\mu\mathbf{x}}, H] = [B_{\mu\mathbf{x}}, H_F] + [B_{\mu\mathbf{x}}, H_Q] \\ &= \nabla^2 A_{\mu\mathbf{x}} + \int [B_{\mu\mathbf{x}}, A^\nu{}_{\mathbf{x}'}]\, j_{\nu\mathbf{x}'}\, d^3x' \\ &= \nabla^2 A_{\mu\mathbf{x}} + 4\pi j_{\mu\mathbf{x}}. \end{aligned} \tag{94}$$

De (93) y (94) resulta

$$\Box A_\mu = 4\pi j_\mu, \tag{95}$$

que concuerda con la teoría de Maxwell y nos dice que la interacción (90) da correctamente la acción de los electrones y positrones sobre el campo.

Para completar la teoría hemos de introducir las condiciones suplementarias (54). Hemos de comprobar que están de acuerdo con las ecuaciones del movimiento. El método empleado en § 77, que consistía en demostrar que las condiciones suplementarias en los distintos instantes en la imagen de Heisenberg eran coherentes, deja de ser aplicable ahora, pues las condiciones cuánticas que relacionan las variables dinámicas en los distintos instantes quedan modificadas por la interacción de forma demasiado complicada para ser desarrolladas. Así pues, lo que haremos será obtener todas las condiciones suplementarias que afectan a las variables dinámicas en un instante y comprobar que son coherentes.

Tenemos nuevamente las ecuaciones (56). Derivando de nuevo respecto a x_0 resulta

$$\Box\partial A_\mu/\partial x_\mu \approx 0. \tag{96}$$

Por otro lado, la ecuación de movimiento (92) para ψ, como en § 68, nos lleva a

$$\partial(\bar{\psi}^\dagger \alpha_\mu \psi)/\partial x_\mu = 0,$$

que es equivalente a

$$\partial j_\mu/\partial x_\mu = 0, \tag{97}$$

pues la diferencia entre $-e\bar{\psi}^\dagger\psi$ y j_0 aunque sea infinita es constante en el tiempo. De (95) resulta que (96) se verifica también como ecuación fuerte. Luego, las ecuaciones (56) son las únicas condiciones suplementarias independientes que afectan a las variables dinámicas en un instante de tiempo. De la primera resulta, como antes, (57), y de la segunda, con ayuda de (95) para $\mu = 0$, obtenemos ahora

$$(A_0{}^r + B_r)^r + 4\pi j_0 \approx 0. \tag{98}$$

Ésta puede escribirse

$$(A_0 + U)^{rr} + 4\pi j_0 \approx 0 \tag{99}$$

o también, según (39),

$$\operatorname{div} \boldsymbol{\mathcal{E}} - 4\pi j_0 \approx 0, \tag{100}$$

que es precisamente una de las ecuaciones de Maxwell.

Puede verse, sin que hagamos los cálculos, que en un mismo instante, para dos puntos x y x'

$$[j_{0\mathbf{x}}, j_{0\mathbf{x}'}] = 0,$$

pues la diferencia entre $-e\bar{\psi}^\dagger\bar{\psi}$ y j_0 aunque sea infinita es constante en el no puede contener derivadas de $\delta(\mathbf{x} - \mathbf{x}')$, y a la vez tiene que ser antisimétrico en $\mathbf{x}$ y $\mathbf{x}'$. Luego, los términos extra $4\pi j_{0\mathbf{x}}$ de la ecuación (98) para los distintos valores de $\mathbf{x}$, comparados con las correspondientes ecuaciones (58), conmutan entre sí y a su vez con todas las variables dinámicas que figuran en (58) y (57). Por consiguiente, dichos términos extra no alterarán la coherencia de (58) y (57), luego, (98) y (57) son coherentes.

Nuestro método de introducir la interacción en la teoría no es relativista, pues la energía de interacción (90) llevaba consigo variables dinámicas en un instante en una referencia de Lorentz particular. Así pues, podemos preguntarnos si la teoría con interacción es relativista o no. Las ecuaciones de campo (92) y (95) así como las condiciones suplementarias (54), son evidentemente relativistas. Queda por saber si las condiciones cuánticas son invariantes bajo las transformaciones de Lorentz.

Conocemos las condiciones cuánticas que relacionan todas nuestras variables dinámicas $A_{\mu\mathbf{x}}$, $B_{\mu\mathbf{x}}$, $\psi_{a\mathbf{x}}$, $\bar{\psi}_{a\mathbf{x}}$ en un instante dado x_0. Como ya hemos dicho, no podemos desarrollar las condiciones cuánticas generales que relacionan las variables dinámicas en dos puntos cualesquiera del espacio-tiempo, porque la interacción las hace demasiado complicadas. Por

tanto, lo que haremos será llevar a cabo una transformación de Lorentz infinitesimal y desarrollar las condiciones cuánticas en el nuevo sistema de referencia. Si podemos demostrar que las condiciones cuánticas en un instante dado son invariantes frente a las transformaciones de Lorentz infinitesimales, quedará demostrado también su invariancia frente a las transformaciones de Lorentz finitas.

Sea x_0^* la coordenada temporal en el nuevo sistema de referencia. Estará relacionada con las coordenadas originales mediante la ecuación

$$x_0^* = x_0 + \varepsilon v_r x_r, \tag{101}$$

siendo ε un número infinitesimal y v_r un vector tridimensional tal que εv_r sea la velocidad relativa entre las dos referencias. Despreciaremos los términos del orden de ε^2

Una cantidad de campo $\varkappa$ en el punto $\mathbf{x}$ y en el instante x_0^* de la nueva referencia tendrá el valor

$$\varkappa(\mathbf{x}, x_0^*) = \varkappa(\mathbf{x}, x_0) + (x_0^* - x_0)\,\partial\varkappa_{\mathbf{x}}/\partial x_0 = \varkappa(\mathbf{x}, x_0) + \varepsilon v_r x)[\varkappa_{\mathbf{x}}, H]. \tag{102}$$

Su P.B. con otra cantidad de campo análoga $\lambda(\mathbf{x}', x_0^*)$ será

$$\begin{aligned}[\varkappa(\mathbf{x}, x_0^*), \lambda(\mathbf{x}', x_0^*)] &= [\varkappa(\mathbf{x}, x_0) + \varepsilon v_r x_r[\varkappa_{\mathbf{x}}, H], \lambda(\mathbf{x}', x_0) + \varepsilon v_s x'_s[\lambda_{\mathbf{x}'}, H]] \\ &= [\varkappa(\mathbf{x}, x_0), \lambda(\mathbf{x}', x_0)] + \varepsilon v_s x'_s[\varkappa_{\mathbf{x}}, [\lambda_{\mathbf{x}'}, H]] + \\ &\qquad + \varepsilon v_r x_r[[\varkappa_{\mathbf{x}}, H], \lambda_{\mathbf{x}'}] \\ &= [\varkappa(\mathbf{x}, x_0), \lambda(\mathbf{x}', x_0)] + \varepsilon v_r(x'_r - x_r)[\varkappa_{\mathbf{x}}, [\lambda_{\mathbf{x}'}, H]] + \\ &\qquad + \varepsilon v_r x_r[[\varkappa_{\mathbf{x}}, \lambda_{\mathbf{x}'}], H].\end{aligned} \tag{103}$$

Si $\varkappa$ y λ son variables ψ o $\bar{\psi}$, nos interesará su anticonmutador y no su P.B. Utilizando la notación (69) del anticonmutador, tenemos

$$\begin{aligned}&[\varkappa(\mathbf{x}, x_0^*), \lambda(\mathbf{x}' x_0^*)]_+ \\ &\quad = [\varkappa(\mathbf{x}, x_0), \lambda(\mathbf{x}', x_0)]_+ + \varepsilon v_r x'_r[\varkappa_{\mathbf{x}}, [\lambda_{\mathbf{x}'}, H]]_+ + \varepsilon v_r x_r[[\varkappa_{\mathbf{x}}, H], \lambda_{\mathbf{x}'}]_+ \\ &\quad = [\varkappa(\mathbf{x}, x_0), \lambda(\mathbf{x}', x_0)]_+ + \varepsilon v_r(x'_r - x_r)[\varkappa_{\mathbf{x}}, [\lambda_{\mathbf{x}'}, H]]_+ + \varepsilon v_r x_r[[\varkappa_{\mathbf{x}}, \lambda_{\mathbf{x}'}]_+, H].\end{aligned} \tag{104}$$

Si $\varkappa$ y λ son dos variables básicas cualesquiera $A_\mu, B_\mu, \psi_a, \bar{\psi}_a$, el P.B. $[\varkappa_{\mathbf{x}}, \lambda_{\mathbf{x}'}]$ o el anticonmutador $[\varkappa_{\mathbf{x}}, \lambda_{\mathbf{x}'}]_+$, según los casos, es un número y, en consecuencia, el último término de (103) o (104) es nulo. Llegamos pues a

$$\begin{aligned}[\varkappa(\mathbf{x}, x_0^*), \lambda(\mathbf{x}', x_0^*)]_\pm &= [\varkappa(\mathbf{x}, x_0), \lambda(\mathbf{x}', x_0)]_\pm + \\ &+ \varepsilon v_r(x_r' - x_r)[\varkappa_{\mathbf{x}}, [\lambda_{\mathbf{x}'}, H_P + H_F]]_\pm + \varepsilon v_r(x'_r - x_r)[\varkappa_{\mathbf{x}}, [\lambda_{\mathbf{x}'}, H_Q]]_\pm,\end{aligned} \tag{105}$$

donde $[\varkappa, \lambda]_\pm$ significa P.B. o anticonmutador, según los casos. A partir de la forma (90) de H_Q vemos que $[\lambda_{\mathbf{x}'}, H_Q]$ sólo puede contener las variables

dinámicas $A_{\mu\mathbf{x}'}$, $\psi_{a\mathbf{x}'}$, $\bar{\psi}_{a\mathbf{x}'}$ no pudiendo contener las derivadas de dichas variables. Por tanto, si $[\varkappa_{\mathbf{x}}, [\lambda_{\mathbf{x}'}, H_Q]]_\pm$ no es nulo, habrá de ser un múltiplo de $\delta(\mathbf{x}-\mathbf{x}')$, y no contendrá términos en las derivadas de $\delta(\mathbf{x}-\mathbf{x}')$. Luego, el último término de (105) es nulo. Resulta, pues, que

$$[\varkappa(\mathbf{x}, x_0^*), \lambda(\mathbf{x}', x_0^*)]_\pm$$

tiene el mismo valor que cuando no hay interacción, y, según lo visto en este caso, es invariante frente a las transformaciones de Lorentz.

Es posible una crítica de la demostración anterior. Hemos desarrollado, en los distintos puntos, expresiones en potencias de ε despreciando ε^2. Por ese procedimiento no podemos calcular $[\varkappa(x), \lambda(x')]_\pm$ para dos puntos cualesquiera del espacio-tiempo x y x' próximos entre sí, para los que $x_\mu - x'_\mu$ sea del orden de ε, pues el resultado de dicho cálculo tiene que ser una función de $(x_\mu - x'_\mu)$ que tiene una singularidad cuando el cuadrivector $x - x'$ está sobre el cono de luz y, en consecuencia, tal función no puede desarrollarse en serie de potencias de las $(x_\mu - x'_\mu)$.

Para que el argumento sea válido hemos de volverlo a formular de modo que se evite el empleo de la función δ. En lugar de

$$[\varkappa(\mathbf{x}, x_0^*), \lambda(\mathbf{x}', x_0^*)]_\pm,$$

hemos de calcular

$$\left[\int a_{\mathbf{x}}\varkappa(\mathbf{x}, x_0^*)\, d^3x, \int b_{\mathbf{x}'}\lambda(\mathbf{x}', x_0^*)\, d^3x'\right]_\pm, \tag{106}$$

donde $a_{\mathbf{x}}$ y $b_{\mathbf{x}}$ son dos funciones continuas arbitrarias de x_1, x_2, x_3. Entonces las cantidades que tenemos que desarrollar en serie de potencias de ε varían todas con continuidad, con un cambio continuo en la dirección del eje de tiempos, quedando así justificados los desarrollos. Las ecuaciones que obtenemos ahora son las de la demostración multiplicadas por $a_{\mathbf{x}}b_{\mathbf{x}'}d^3xd^3x'$ e integradas. Llegamos a la misma conclusión de antes: los P.B. o los anticonmutadoras tienen los mismos valores que en ausencia de interacción.

Podríamos ver que la razón de que la integración no modifique las condiciones cuánticas es su simplicidad, es decir, el hecho de que sólo sea función de las variables dinámicas básicas y no de sus derivadas. Los P.B. y anticonmutadores tienen los mismos valores que en ausencia de interacción, siempre que se refieran a variables de dos puntos del espacio-tiempo consideradas en el mismo instante por algún observador. Esto significa que los dos puntos han de estar cada uno fuera del cono de luz del otro y únicamente se pueden aproximar a través de líneas que vayan por fuera del cono de luz.

80. *Variables físicas*

Todo ket $|P\rangle$ que represente un estado físico ha de verificar las condiciones suplementarias

$$(B_0 + A_r^r)|P\rangle = 0, \qquad (\operatorname{div} \boldsymbol{\mathcal{E}} - 4\pi j_0)|P\rangle = 0. \tag{107}$$

Una variable es física, si al aplicarla a cualquier ket que verifique dichas condiciones da como resultado otro ket que también las verifica. Ello exige que conmute con las cantidades

$$B_0 + A_r^r, \qquad \operatorname{div} \boldsymbol{\mathcal{E}} - 4\pi j_0. \tag{108}$$

Veamos qué variables dinámicas sencillas gozan de esta propiedad.

Las variables de campo transversales $\mathscr{A}_r$, $\mathscr{B}_r$ conmutan con las cantitidades (108), y en consecuencia son variables físicas. La variable ψ_a conmuta con la primera de las cantidades (108) pero no con la segunda, luego, no es una variable física. Tenemos

$$\begin{aligned} i\hbar[\psi_{a\mathbf{x}}, \bar{\psi}_{b\mathbf{x}'} \cdot \psi_{b\mathbf{x}'}] &= (\psi_{a\mathbf{x}} \bar{\psi}_{b\mathbf{x}'} + \bar{\psi}_{b\mathbf{x}'} \psi_{a\mathbf{x}})\psi_{b\mathbf{x}'} \\ &= \delta_{ab}\ \delta(\mathbf{x} - \mathbf{x}')\bar{\psi}_{b\mathbf{x}'} = \psi_{a\mathbf{x}}\, \delta(\mathbf{x} - \mathbf{x}'). \end{aligned}$$

Luego,

$$[\psi_{a\mathbf{x}}, j_{0\mathbf{x}}] = ie/\hbar \cdot \psi_{a\mathbf{x}}\, \delta(\mathbf{x} - \mathbf{x}'). \tag{109}$$

A partir de (42)

$$[e^{ieV_{\mathbf{x}}/\hbar}, \operatorname{div} \boldsymbol{\mathcal{E}}_{\mathbf{x}'}] = 4\pi ie/\hbar \cdot e^{ieV_{\mathbf{x}}/\hbar} \delta(\mathbf{x} - \mathbf{x}').$$

Por tanto,

$$\begin{aligned} [e^{ieV_{\mathbf{x}}/\hbar}\psi_{a\mathbf{x}}, \operatorname{div} \boldsymbol{\mathcal{E}}_{\mathbf{x}'} - 4\pi j_{0\mathbf{x}'}] &= [e^{ieV_{\mathbf{x}}/\hbar}, \operatorname{div} \boldsymbol{\mathcal{E}}_{\mathbf{x}'}]\psi_{a\mathbf{x}} - 4\pi e^{ieV_{\mathbf{x}}/\hbar}[\psi_{a\mathbf{x}}, j_{0\mathbf{x}'}] \\ &= 0. \end{aligned}$$

Luego, si hacemos

$$\psi^*_{a\mathbf{x}} = {}^{ieV_{\mathbf{x}}/\hbar}\psi_{a\mathbf{x}}, \tag{110}$$

$\psi^*_{a\mathbf{x}}$ conmuta con las dos expresiones (108) y es una variable física. Análogamente también es variable física $\bar{\psi}^*_{a\mathbf{x}}$. Las variables $\mathscr{A}_r$, $\mathscr{B}_r$, ψ^*_a, $\bar{\psi}^*_a$ son las únicas variables físicas independientes, además de las propias cantidades (108).

Tenemos

$$j_0 = -\tfrac{1}{2}e(\bar{\psi}^{*\dagger}\psi^* - \psi^{*\dagger}\bar{\psi}^*), \qquad j_r = -e\bar{\psi}^{*\dagger}\alpha_r\psi^*. \tag{111}$$

Luego, la densidad de carga y la corriente son variables físicas. También es fácil ver que $\boldsymbol{\mathcal{E}}$ y $\boldsymbol{\mathcal{H}}$ son variables físicas, como sucedía cuando no existían electrones ni positrones. Todas las variables que no quedan modificadas por la arbitrariedad que existe en la determinación de los potenciales electromagnéticos en la teoría de Maxwell, son variables físicas.

El operador $\psi_{a\mathbf{x}}$ representa la creación de un positrón o la aniquilación de un electrón en el lugar $\mathbf{x}$. Veamos cuál es el significado físico del operador $\psi^*_{a\mathbf{x}}$. Según (44),

$$i\hbar[e^{ieV_{\mathbf{x}}/\hbar}, \mathcal{E}_{r\mathbf{x}'}] = ee^{ieV_{\mathbf{x}}/\hbar}(x_r - x'_r)|\mathbf{x} - \mathbf{x}'|^{-3},$$

y por tanto,

$$i\hbar[\psi^*_{a\mathbf{x}}, \mathcal{E}_{r\mathbf{x}'}] = e\psi^*_{a\mathbf{x}}(x_r - x'_r)|\mathbf{x} - \mathbf{x}'|^{-3}$$

o sea,

$$\mathcal{E}_{r\mathbf{x}'}\psi^*_{a\mathbf{x}} = \psi^*_{a\mathbf{x}}\{\mathcal{E}_{r\mathbf{x}'} + e(x'_r - x_r)|\mathbf{x}' - \mathbf{x}|^{-3}\}. \tag{112}$$

Consideremos un estado $|P\rangle$ en el cual $\mathcal{E}_r$ tenga ciertamente el valor c_r en el punto $\mathbf{x}'$, o sea,

$$\mathcal{E}_{r\mathbf{x}'}|P\rangle = c_r|P\rangle.$$

Entonces según (112)

$$\mathcal{E}_{r\mathbf{x}'}\psi^*_{a\mathbf{x}}|P\rangle = \{c_r + e(x'_r - x_r)|\mathbf{x}' - \mathbf{x}|^{-3}\}\psi^*_{a\mathbf{x}}|P\rangle,$$

luego, en el estado $\psi^*_{a\mathbf{x}}|P\rangle$, $\mathcal{E}_r$ en el punto $\mathbf{x}'$ tiene ciertamente el valor

$$c_r + e(x'_r - x_r)|\mathbf{x}' - \mathbf{x}|^{-3}.$$

Esto significa que el operador $\psi^*_{a\mathbf{x}}$, además de crear un positrón o aniquilar un electrón en el punto $\mathbf{x}$, aumenta el campo eléctrico en el punto $\mathbf{x}'$ en la cantidad $e(x'_r - x_r)|\mathbf{x}' - \mathbf{x}|^{-3}$, que coincide precisamente con el campo de Coulomb clásico en el punto $\mathbf{x}'$ creado por un positrón de carga e en el punto $\mathbf{x}$. Luego, el operador $\psi^*_{a\mathbf{x}}$ crea un positrón en el punto $\mathbf{x}$ *junto con su campo coulombiano*, o bien aniquila un electrón en el lugar $\mathbf{x}$ *junto con su campo coulombiano.*

Para electrones y positrones en interacción con el campo electromagnético, más que las variables ψ, $\bar{\psi}$, son las variables ψ^*, $\bar{\psi}^*$ las que corresponden al proceso físico de creación o aniquilación de electrones y positrones, pues dichos procesos siempre deben ir acompañados por el cambio correspondiente de Coulomb en el campo eléctrico alrededor del punto en que ha sido creada o aniquilada la partícula. Puede verse fácilmente que las variables $\psi^*_{a\mathbf{x}}, \bar{\psi}^*_{a\mathbf{x}}$ verifican las mismas relaciones de anticonmutación (70) que las variables sin estrella. Al pasar a la representación de momentos, las cantidades importantes no serán las variables no físicas $\psi_{\mathbf{p}}$ definidas por (67), sino las variables físicas $\psi^*_{\mathbf{p}}$ definidas por

$$\psi^*_{\mathbf{x}} = h^{-\frac{3}{2}} \int e^{i(\mathbf{xp})/\hbar}\psi^*_{\mathbf{p}}\, d^3p, \qquad \psi^*_{\mathbf{p}} = h^{-\frac{3}{2}} \int e^{-i(\mathbf{xp})/\hbar}\psi^*_{\mathbf{x}}\, d^3x. \tag{113}$$

Ahora hemos de sustituir (68) por

$$\xi^*_{\mathbf{p}} = \frac{1}{2}\left\{1 + \frac{\alpha_r p_r + \alpha_m m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}\psi^*_{\mathbf{p}} \quad , \quad \zeta^*_{\mathbf{p}} = \frac{1}{2}\left\{1 - \frac{\alpha_r p_r + \alpha_m m}{(\mathbf{p}^2 + m^2)^{\frac{1}{2}}}\right\}\psi^*_{\mathbf{p}}$$

y tomar $\xi^*_{\mathbf{p}}$ para representar la aniquilación de un electrón de momento $\mathbf{p}$, $\bar{\xi}^*_{\mathbf{p}}$ para la creación de un electrón de momento $\mathbf{p}$, $\zeta^*_{\mathbf{p}}$ para la creación de un positrón de momento $-\mathbf{p}$ y $\bar{\zeta}^*_{\mathbf{p}}$ para la aniquilación de un positrón de momento $-\mathbf{p}$. Las variables $\psi^*_{\mathbf{p}}$, $\bar{\psi}^*_{\mathbf{p}}$, $\xi^*_{\mathbf{p}}$, $\bar{\xi}^*_{\mathbf{p}}$, $\zeta^*_{\mathbf{p}}$, $\bar{\zeta}^*_{\mathbf{p}}$, verificarán las mismas relaciones de anticonmutación que las correspondientes variables sin estrella.

Podemos expresar el hamiltoniano enteramente en función de variables físicas. Tenemos

$$\psi^{*r} = e^{ieV/\hbar}(\psi^r + ie/\hbar \cdot V^r\psi).$$

Luego,

$$H_P + H_Q = \int \{\bar{\psi}^\dagger\alpha^r[-i\hbar\psi^r - e(\mathscr{A}^r - V^r)\psi] + \bar{\psi}^\dagger\alpha_m m\psi + A^0 j_0\}\, d^3x$$
$$= \int \{\bar{\psi}^{*\dagger}\alpha_r(-i\hbar\psi^{*r} - e\mathscr{A}^r\psi^*) + \bar{\psi}^{*\dagger}\alpha_m m\psi^* + A^0 j_0\}\, d^3x.$$

El último término del integrando ha de combinarse con H_{FL}. De (49) y (57)

$$H_{FL} \approx -(8\pi)^{-1} \int (U - A_0)(U + A_0)^{rr}\, d^3x$$
$$\approx \tfrac{1}{2} \int (U - A_0) j_0\, d^3x$$

con ayuda de (99). Luego,

$$H_{FL} + \int A^0 j_0\, d^3x \approx \tfrac{1}{2} \int (U + A_0) j_0\, d^3x.$$

Integrando (99) con ayuda de la fórmula (72) de § 38 resulta

$$A_{0\mathbf{x}} + U_{\mathbf{x}} \approx \int \frac{j_{0\mathbf{x}'}}{|\mathbf{x} - \mathbf{x}'|}\, d^3x',$$

y por tanto,

$$H_{FL} + \int A^0 j_0\, d^3x \approx \frac{1}{2} \int\int \frac{j_{0\mathbf{x}} j_{0\mathbf{x}'}}{|\mathbf{x} - \mathbf{x}'|}\, d^3x d^3x'.$$

Luego, se tiene

$$H \approx H^*$$

con

$$H^* = \int \{\bar{\psi}^{*\dagger}\alpha_r(-i\hbar\psi^{*r} - e\mathscr{A}^r\psi^*) + \bar{\psi}^{*\dagger}\alpha_m m\psi^*\}\, d^3x +$$
$$+ H_{FT} + \frac{1}{2}\int\int \frac{j_{0\mathbf{x}} j_{0\mathbf{x}'}}{|\mathbf{x} - \mathbf{x}'|}\, d^3x d^3x'. \quad (114)$$

En la ecuación de Schrödinger de un ket $|P\rangle$ que representa un estado físico, podemos emplear H^* en lugar de H, pues para dicho ket se tiene

$$i\hbar d|P\rangle/dx_0 = H|P\rangle = H^*|P\rangle. \quad (115)$$

H^* sólo contiene variables físicas. Las variables de campo longitudinales no aparecen. En su lugar aparece el último término de (114), que es precisamente la energía de interacción de Coulomb de todas las cargas presentes. Es bastante extraño que aparezca un término de este tipo en una teoría relativista, pues representa la energía asociada con la propagación

instantánea de fuerzas. Dicho término aparece como resultado de habernos separado de la imagen de Heisenberg, que es donde se pone de manifiesto la invariancia de la teoría.

Podríamos haber establecido una representación tomando como ket standard el producto del ket standard $|0_F\rangle$ del campo electromagnético aislado, dado por (61) y (62), por el ket standard $|0_P\rangle$ de los electrones y positrones aislados, dado por (74). Sin embargo, esta representación no sería adecuada, pues su ket standard no verificaría la segunda de las condiciones suplementarias de (107).

Otra representación más conveniente se obtiene tomando un ket standard $|0_Q\rangle$ que verifique

$$(B_0 + A_r^r)|0_Q\rangle = 0, \qquad (\operatorname{div} \mathcal{E} - 4\pi j_0)|0_Q\rangle = 0, \tag{116}$$

$$\mathcal{A}_{r\mathbf{k}}|0_Q\rangle = 0, \quad \xi^*_{a\mathbf{p}}|0_Q\rangle = 0, \quad \zeta^*_{a\mathbf{p}}|0_Q\rangle = 0. \tag{117}$$

Estas condiciones son coherentes, pues los operadores que actúan sobre $|0_Q\rangle$ en ellas conmutan o anticonmutan entre sí, y además son suficientes para fijar completamente $|0_Q\rangle$ salvo en un factor numérico, pues hay el mismo número de ecuaciones que las que se necesitan para fijar $|0_F\rangle|0_P\rangle$. Las condiciones (116) muestran que $|0_Q\rangle$ verifica las condiciones suplementarias y, por tanto, representa un estado físico. Las condiciones (117) nos dicen que $|0_Q\rangle$ representa un estado en el que no existen fotones, electrones ni positrones.

Cualquier ket $|P\rangle$ que verifique las condiciones suplementarias (107) y que, en consecuencia, represente un estado físico puede expresarse en forma de una variable física aplicada al ket $|0_Q\rangle$. Las únicas variables físicas independientes que no dan resultados nulos al ser aplicadas a $|0_Q\rangle$ son $\mathcal{A}_{r\mathbf{k}}$, $\xi^*_{a\mathbf{p}}$, $\zeta^*_{a\mathbf{p}}$. Luego,

$$|P\rangle = \Psi(\mathcal{A}_{r\mathbf{k}}, \xi^*_{a\mathbf{p}}, \zeta^*_{a\mathbf{p}})|0_Q\rangle. \tag{118}$$

Así pues, $|P\rangle$ está representado por un funcional de onda Ψ de las variables $\mathcal{A}_{r\mathbf{k}}$, $\xi^*_{a\mathbf{p}}$, $\zeta^*_{a\mathbf{p}}$. Es una serie de potencias en dichas variables, cuyos distintos términos corresponden a la existencia de los distintos números de fotones, electrones y positrones con sus campos coulombianos.

Empleando la representación (118) y el hamiltoniano H^*, podemos no tener en cuenta las condiciones (116), ya que no jugarán ningún papel en la resolución de la ecuación de Schrödinger (115). Las variables longitudinales ya no vuelven a aparecer en la teoría.

81. *Dificultades de la teoría*

El ket standard $|0_Q\rangle$ representa un estado en el que no existen fotones, electrones ni positrones. Podríamos inclinarnos a pensar que dicho estado representa el vacío perfecto, pero no puede ser así pues no es estacionario. Para que fuese estacionario necesitaríamos que

$$H^*|0_Q\rangle = C|0_Q\rangle$$

siendo C un número. Pero H^* contiene los términos

$$-e\int \bar{\psi}^{*\dagger}\alpha_r \mathscr{A}^r\psi^* \, d^3x + \tfrac{1}{2}\int\int \frac{j_{0\mathbf{x}}j_{0\mathbf{x}'}}{|\mathbf{x}-\mathbf{x}'|}\, d^3x d^3x', \tag{119}$$

que al aplicarlos al ket $|0_Q\rangle$ no dan factores numéricos, y en consecuencia, destruyen el carácter estacionario de $|0_Q\rangle$.

Dichos términos pueden ser considerados como una perturbación que da lugar a una probabilidad de que el estado $|0_Q\rangle$ pase a otro estado, de acuerdo con la teoría de § 44. El primero de dichos términos descompuesto en componentes de Fourier contiene una parte

$$-e(\alpha_r)_{ab}\int\int \mathscr{A}^r{}_{\mathbf{k}}\, \xi^*_{a\mathbf{p}}\, \zeta^*_{b\mathbf{p}+\mathbf{k}\hbar}\, d^3k d^3p, \tag{120}$$

que da lugar a transiciones en las que se emite un fotón y en las que simultáneamente se crea un par electrón-positrón. Tras un corto intervalo de tiempo, la probabilidad de transición es proporcional al cuadrado de la longitud del ket que se obtiene aplicando (120) al ket inicial $|0_Q\rangle$, y vale

$$\begin{aligned} &e^2(\bar{\alpha}_r)_{ab}(\alpha_s)_{cd}\times \\ &\quad\times\int\int\int\int \langle 0_Q|\bar{\zeta}^*_{a\,\mathbf{p}+\mathbf{k}\hbar}\,\xi^*_{b\mathbf{p}}\;\mathscr{A}^r{}_{\mathbf{k}}\mathscr{A}^s{}_{\mathbf{k}'}\xi^*_{c\mathbf{p}'}\,\zeta^*_{d\,\mathbf{p}'+\mathbf{k}'\hbar}\,|0_Q\rangle d^3k d^3p d^3k' d^3p' \\ &= e^2(\bar{\alpha}_r)_{ab}(\alpha_s)_{cd}\int\int\int\int \langle 0_Q|i\hbar[\mathscr{A}^r{}_{\mathbf{k}}, \mathscr{A}^s{}_{\mathbf{k}'}]\times \\ &\qquad\times[\xi^*_{b\mathbf{p}'}, \xi^*_{c\mathbf{p}'}]_+[\bar{\zeta}^*_{a\,\mathbf{p}+\mathbf{k}\hbar}, \zeta^*_{d\,\mathbf{p}'+\mathbf{k}''\hbar}]_+|0_Q\rangle d^3k d^3p d^3k' d^3p'. \end{aligned}$$

Empleando los valores de los P.B. y anticonmutadores dados por (4), (16), (72), (73) se obtiene un integrando que depende de las variables $\mathbf{k}$, $\mathbf{k}'$ según la ley $|\mathbf{k}|^{-1}\delta(\mathbf{k}-\mathbf{k}')$ para valores grandes de $\mathbf{k}$ y $\mathbf{k}'$. Con ello resulta una integral divergente, y así la probabilidad de transición es infinita.

El segundo término de (119), descompuesto en componentes de Fourier, contiene términos como $\xi^*_{\mathbf{p}}\,\xi^*_{\mathbf{p}'}\,\zeta^*_{\mathbf{p}''}\,\zeta^*_{\mathbf{p}+\mathbf{p}'-\mathbf{p}''}$ que dan lugar a transiciones en las que se crean simultáneamente dos pares electrón-positrón. Podemos calcular igualmente esta probabilidad de transición, obteniéndose también un valor infinito.

Según estos cálculos resulta que el estado 0_Q no es estacionario ni

siquiera aproximadamente. Una dificultad más seria es el hecho de que no se pueda utilizar el método de perturbaciones para tratar la ecuación de ondas cuando conduce a probabilidades de transición del estado inicial que se considere. En cambio, no se conoce ningún otro método capaz de sustituirlo y, por tanto, no se conoce ninguna solución de la ecuación de onda de la electrodinámica cuántica.

El origen de este conflicto reside en las fluctuaciones de la densidad del mar de electrones de energía negativa, fluctuaciones que se presentan siempre en la teoría de electrones y positrones en ausencia de campo electromagnético. La densidad de electrones en estados tanto de energía positiva como de energía negativa es

$$\rho_{\mathbf{x}} = \psi_{\mathbf{x}}^{\dagger}\psi_{\mathbf{x}} = (\xi_{\mathbf{x}}^{\dagger} + \zeta_{\mathbf{x}}^{\dagger})(\xi_{\mathbf{x}} + \zeta_{\mathbf{x}}).$$

Por tanto, según (74)

$$\rho_{\mathbf{x}}|0_P\rangle = \xi_{\mathbf{x}}^{\dagger}\xi_{\mathbf{x}}|0_P\rangle + (\zeta_{\mathbf{x}}^{\dagger}\zeta_{\mathbf{x}} + (\zeta_{\mathbf{x}}^{\dagger}\zeta_{\mathbf{x}})|0_P\rangle. \tag{121}$$

Las relaciones de anticonmutación (73) para las ζ y las ξ nos dicen que

$$\zeta_{\mathbf{x}}^{\dagger}\zeta_{\mathbf{x}} + \zeta_{\mathbf{x}}^{\dagger}\zeta_{\mathbf{x}} \tag{122}$$

es un número, aunque sea infinito, y por tanto el segundo término del segundo miembro de (121) es un múltiplo numérico de $|0_P\rangle$. En cambio, el primero no puede ser nunca un múltiplo numérico de $|0_P\rangle$, y en consecuencia, $|0_P\rangle$ no es un autoket de $\rho_{\mathbf{x}}$. Esto implica que la densidad del mar de electrones de energía negativa en el vacío no es constante, sino que está sometida a fluctuaciones continuas.

Siguiendo la idea de § 73, a fin de obtener una cantidad físicamente importante — la densidad de electrones de energía positiva menos la densidad de positrones — restaremos de la densidad de todos los electrones la densidad de la distribución de vacío. La cantidad que así se obtiene es la densidad de carga que debe emplearse en la ecuación de Maxwell (58) de § 73. En una teoría exacta de la electrodinámica cuántica, la cantidad que restamos ha de ser un operador bien definido, que ha de poder ser introducido en las ecuaciones. Lo mejor que podemos hacer es tomar para dicho operador la parte constante (122) de la densidad de la distribución de vacío dejando de restar la parte que da lugar a fluctuaciones. Con ello resulta

$$j_{0\mathbf{x}} = -e\{\psi_{\mathbf{x}}^{\dagger}\psi_{\mathbf{x}} - (\zeta_{\mathbf{x}}^{\dagger}\zeta_{\mathbf{x}} + \zeta_{\mathbf{x}}^{\dagger}\zeta_{\mathbf{x}})\}. \tag{123}$$

Con ayuda de la relación

$$\xi_{\mathbf{x}}^{\dagger}\xi_{\mathbf{x}} + \xi_{\mathbf{x}}^{\dagger}\xi_{\mathbf{x}} = \zeta_{\mathbf{x}}^{\dagger}\zeta_{\mathbf{x}} + \zeta_{\mathbf{x}}^{\dagger}\zeta_{\mathbf{x}},$$

que se obtiene a partir de las relaciones de anticonmutación (72) y (73) haciendo $a = b$ y sumando, resulta inmediatamente que (123) coincide con la expresión (77) que teníamos para j_0.

Pero como no hemos restado la parte de la densidad de la distribución de vacío que da lugar a fluctuaciones, ésta continúa formando parte de la densidad de carga que se aplica al ket de vacío, es decir,

$$j_{0\mathbf{x}}|0_P\rangle = -e\xi^\dagger_{\mathbf{x}}\zeta_{\mathbf{x}}|0_P\rangle.$$

Luego, la densidad de carga del vacío no es cero. Llegamos al resultado aparentemente paradójico de que, pese a que el estado vacío ha sido definido como un estado en el que no existe ningún electrón ni ningún positrón, sigue teniendo densidad de carga no nula. Este resultado es una consecuencia necesaria de no haber podido restar las fluctuaciones del mar de electrones de energía negativa, incluso en el caso en que no hay interacción con el campo electromagnético.

Existe también alguna dificultad en relación con el término constante (122) de la densidad de distribución de vacío. Por una consideración de simetría, es de esperar que no aparezca ningún término constante en la corriente producida por el movimiento de los electrones de la distribución de vacío, con lo que tendremos una densidad de carga a la que no está asociada ninguna corriente. Esto determina un sistema de referencia de Lorentz privilegiado, cosa que en una teoría relativista no debería ocurrir.

El punto de apoyo para salvar esta incoherencia reside en que la corriente producida por el movimiento de los electrones de la distribución de vacío, en realidad, es una diferencia entre dos infinitos, y por consiguiente, no es una cantidad bien definida. Dicho término constante puede ser obtenido por un paso al límite, llevado a cabo en un sistema de referencia de Lorentz particular, obteniendo así un valor particular; pero dicho valor no debería ser más privilegiado que otros. La ambigüedad en el término constante de la corriente no representa una dificultad seria para la teoría, ya que podemos evitar dicha ambigüedad definiendo de nuevo j_r por

$$j_{r\mathbf{x}} = -\tfrac{1}{2}e(\psi^\dagger_{\mathbf{x}}\alpha_r\psi_{\mathbf{x}} - \psi^\dagger_{\mathbf{x}}\alpha^\dagger_r\psi_{\mathbf{x}})$$

en lugar de por (78). Con ello queda en la misma forma que la expresión (77) de j_0, y nos permite escribir la ecuación relativista

$$j_\mu = -\tfrac{1}{2}e(\psi^\dagger\alpha_\mu\psi - \psi^\dagger\alpha^\dagger_\mu\psi). \tag{124}$$

La dificultad de las fluctuaciones es mucho más seria. Al poner en interacción la densidad de carga fluctuante del vacío con el campo electromagnético ésta da lugar a la aparición de un campo eléctrico fluctuante, según se deduce de la ecuación de Maxwell (100), que actúa sobre el mar de electrones de energía negativa dando lugar a la creación de pares

electrón-positrón. Si intentamos calcular esta acción nos encontramos con integrales divergentes, lo que impide obtener una solución de la ecuación de onda con interacción.

Ya habían aparecido antes en la teoría algunas cantidades infinitas, pero ninguna de ellas constituía una dificultad seria, pues no jugaban ningún papel importante en las ecuaciones; podíamos evitarlas volviendo a definir las cantidades de forma conveniente. Así teníamos el infinito del punto cero de la energía del campo electromagnético transversal H_{FT} dada por (47), que pudimos evitar empleando (65). Nos encontrábamos también con el infinito del punto cero de la energía de los electrones y positrones H_P, dada por (80), y con el término infinito que aparecía explícitamente en la forma (82) de H_P. Dicho infinito nos fue posible evitarlo considerando H_P en la forma (81). Apareció también la parte constante infinita de la densidad de electrones del vacío (122), que pudimos eliminar de la teoría empleando la fórmula (124) para j_μ. En cambio, no existe ninguna posibilidad de volver a formular la teoría que nos permita evitar los infinitos de las fluctuaciones.

Se ha conseguido establecer ciertas reglas que permiten eliminar de un modo coherente los infinitos que se derivan de las fluctuaciones, habiéndose obtenido con ello una teoría capaz de ser desarrollada, y a partir de la cual se pueden obtener resultados que pueden ser confrontados con los experimentos. En la confrontación se ha encontrado buena concordancia, lo que indica que las reglas tienen cierta validez. Pero dichas reglas sólo se pueden aplicar a problemas particulares, que por lo general son problemas de colisión, y no encajan en los fundamentos lógicos de la mecánica cuántica. Por tanto, no pueden ser consideradas como una solución satisfactoria de las dificultades.

Parece ser que hemos seguido hasta donde es posible el desarrollo lógico de las ideas de la mecánica cuántica tal y cómo se conocen hoy en día. Teniendo en cuenta que las dificultades son de carácter muy profundo, únicamente pueden ser superadas por un cambio drástico de los fundamentos de la teoría, probablemente tan drástico como el paso de la teoría de las órbitas de Bohr a la mecánica cuántica actual.

INDICE ALFABETICO

acción, 140
adjunto, 38
amplitud de probabilidad, 84
— — relativa, 84
anomalía magnética del spin, 179
anticonmutador, 310
anticonmutar, 162
antilineal, 33
armónicos esféricos, 168
auto-, 41
autoadjunto, 39
autoestado común, 60
autofunción, 129
Bohr, condición de las frecuencias de, 129, 190
bosón, 224
bra, 31
bras básicos, 64
— ortogonales, 33
campo central, 165
— longitudinal, 293
— transversal, 293
carácter de un grupo, 229
causalidad, 18
clase de una permutación, 226
coeficiente de probabilidad, 193
complejo conjugado, 22
condición de contorno, 168
condiciones cuánticas, 95
— suplementarias, 304
conjunto completo de bras, 64
— — de estados, 48
— — de observables que conmutan, 67
— exclusivo de estados, 229
conmutación, 36
constante del movimiento, 127
contravariante, 269
coordenadas canónicas conjugadas, 96
corriente de probabilidad, 275
Coulomb, energía de interacción, 322
covariante, 269
δ_{rs}, 73
de Broglie, ondas de, 132
densidad de probabilidad, 273
dependiente, 28, 29
diagonal en la representación, 84
— respecto a un observable, 88
dispersor, 199
e, 169
ecuación débil, 306
— de ondas, 123
— fuerte, 306
Einstein, efecto fotoeléctrico, 21
elemento de matriz, 78
— diagonal, 79
emisión estimulada, 190, 253
energía longitudinal, 303, 307
— propia, 192
— transversa, 303, 308
espacio de las fases, 144
estadística de Bose, 224
— de Fermi, 224
estado, 24
— antisimétrico, 222
— de absorción, 201
— de movimiento, 25
— de polarización, 18
— de traslación, 21
— estacionario, 128
— fundamental, 142
— intermedio, 188
— simétrico, 223
estados ligados, 169
— ortogonales, 33, 47
factor de fase, 34
fermión, 224
Fock, representación de, 151
funcional, 308
función δ, 69
función Δ, 296
función bien ordenada, 142
— de onda, 91
— — dependiente del tiempo, 123
— de peso, 77
— de transformación, 86
— impropia, 69
Gibbs, conjunto de, 143
Green, teorema de, 205
h, $\hbar$, 88
hamiltoniano, 125, 126
Hamilton-Jacobi, ecuación de, 134
Heisenberg, imagen de, 124
— representación de, 129
— variable dinámica de, 125
Hilbert, espacio de, 51
huecos, 267
imaginario conjugado, 33
independiente, 28, 39

ket, 28
— antisimétrico, 222
— básico, 64
— simétrico, 223
— ortogonal, 33
— standard, 90
Kramer-Heisenberg, fórmula de dispersión de, 263
lagrangiano, 140
Landé, fórmula de, 198
ley de combinación de la espectroscopía, 11, 129
leyes de conservación, 127, 155
longitud de un bra o de un ket, 33
matriz, 78, 79
— diagonal, 80, 81, 82
— hermítica, 79, 80.
— unidad, 79, 80.
Maxwell, ecuaciones de, 306, 317
mecánica ondulatoria, 27.
módulo del vector momento angular, 159
momento angular, 152
— — de spin, 155, 282
— — orbital, 155, 161
— magnético del electrón, 178, 281
— radial, 165
multiplete, 196, 237
normalización, 33
notación de la barra encima, 32
observable, 48, 304
— tiene un valor, 58
— tiene un valor medio, 57
observaciones compatibles, 63
operador de antisimetrización, 263
— de rotación, 154
— de simetrización, 239
— de traslación, 113
— lineal, 35
— — complejo conjugado, 39
— — real, 39
oscilador, 148, 242
paquete de ondas, 109, 133
Pauli, principio de exclusión de, 225
P. B., 96
permutación, 222, 225
— equivalente, 226
— idéntica, 225
— impar, 222
— par, 222
— recíproca, 226
pertenece al autovalor, 42
Planck, constante de, 98
Poisson, corchete de, 96
positrón, 289
principio de exclusión, 225
— de incertidumbre, 109
probabilidad de que un observable tenga un valor, 58
raíz cuadrada de un observable, 56
— — positiva, 56
recíproco de un observable, 55
regla de selección, 172
relación de conmutación, 95
representación, 64
— de momentos, 106
— ortogonal, 65
— simétrica, 222
representante, 64, 78
Schrödinger, ecuación de ondas de, 123
— imagen de, 123
— representación de, 105
— variable dinámica de, 125
segunda cuantificación, 244, 265
semiamplitud de la línea de absorción, 218
simetría esférica, 156
sistema degenerado, 184
— no degenerado, 184
Sommerfeld, fórmula de, 288
spin del electrón, 161, 282
suma de momentos angulares, 159
superposición de estados, 24
teorema de ortogonalidad, 44
— de reciprocidad, 87
transformación de contacto, 116
unitario, 116
variable física, 304-305
— orbital, 234
vector dual, 31
velocidad de grupo, 132